Программная система «MABYM»

(MOOD ANALYSIS BY MESSAGES -АНАЛИЗ настроения по сообщениям)

Анализ предметной области

Основная деятельность приложения «MABYM» — это анализ настроения пользователей по сообщениям, с помощью которого можно будет предположить возможный исход матча. Также пользователь может просмотреть самые популярные сообщения(значимые). Приложение анализирует сообщения по заготовленному ранее словарю.

Функционал

* + - 1. предоставлять возможность изменять источники данных для сбора сообщений.

2. сохранять проанализированные сообщения и успешные запросы.

3. Необходимо определять эмоциональную окраску текста на основе выборки сообщений социальной сети «ВКонтакте» по трем категориям: позитивная, нейтральная и негативная.

4. Точность определения тональности не менее 70 %.

5. Система анализа должна быть реализована как web-приложение и быть доступна с любого устройства, имеющего выход в сеть Интернет.

Анализ аналогов

* Один из аналогов - «Brand Analytics». Это система аналитики социальных медиа. Она собирает информацию о нужном бренде и анализирует собранные данные. Может определить отношение к тому или иному бренду с помощью постов и блогов в социальных сетях, а также комментариев к ним. Кроме поиска по ключевым словам, Brand Analytics собирает и анализирует сообщения по чекинам и гео-меткам, избранным авторам и группам в соцсетях.



* «Youscan» - продукт для мониторинга и анализа упоминаний в социальных медиа. Данное решение мониторит все упоминания бренда в соцсетях, блогах и форумах.



СРАВНЕНИЕ МЕТОДОВ АНАЛИЗА ТОНАЛЬНОСТИ ТЕКСТОВ



Определение тональности текстовых сообщений способно выявить эмоционально окрашенную лексику и проанализировать оценку автора по отношению к объектам, речь о которых идет в тексте.

В данной системе будет использован метод анализа тональности текста со словарем. Этот метод основан на поиске эмотивной лексики в тексте по заранее составленным тональным словарям и правилам с применением лингвистического анализа. По совокупности найденной эмотивной лексики текст может быть оценен по шкале, содержащей количество негативной и позитивной лексики. Данный метод может использовать как списки правил, подставляемые в регулярные выражения, так и специальные правила соединения тональной лексики внутри предложения. Чтобы проанализировать текст, можно воспользоваться следующим алгоритмом: сначала каждому слову в тексте присвоить его значение тональности из словаря (если оно присутствует в словаре), а затем вычислить общую тональность всего текста путём суммирования значения тональностей каждого отдельного предложения.

Диаграмма прецедентов

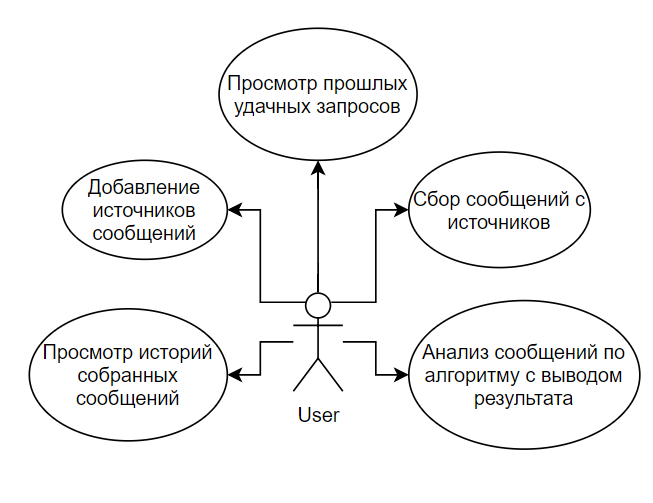


Рисунок 1. Диаграмма прецедентов

ДИАГРАММА КЛАССОВ

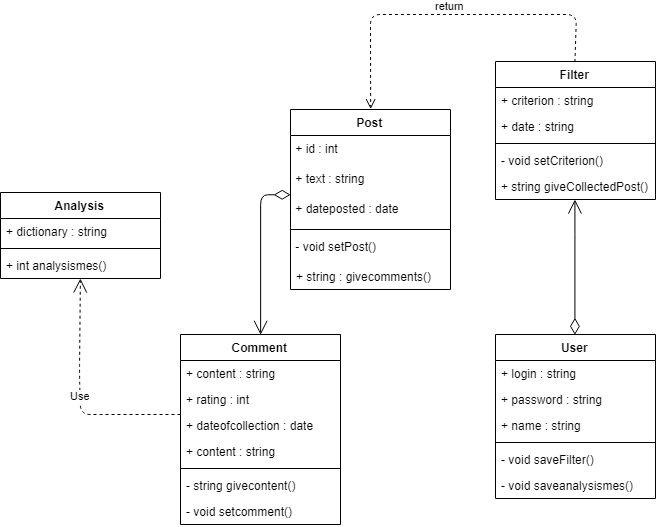


Рисунок 2. Диаграмма классов.

МОДЕЛЬ IDEF0

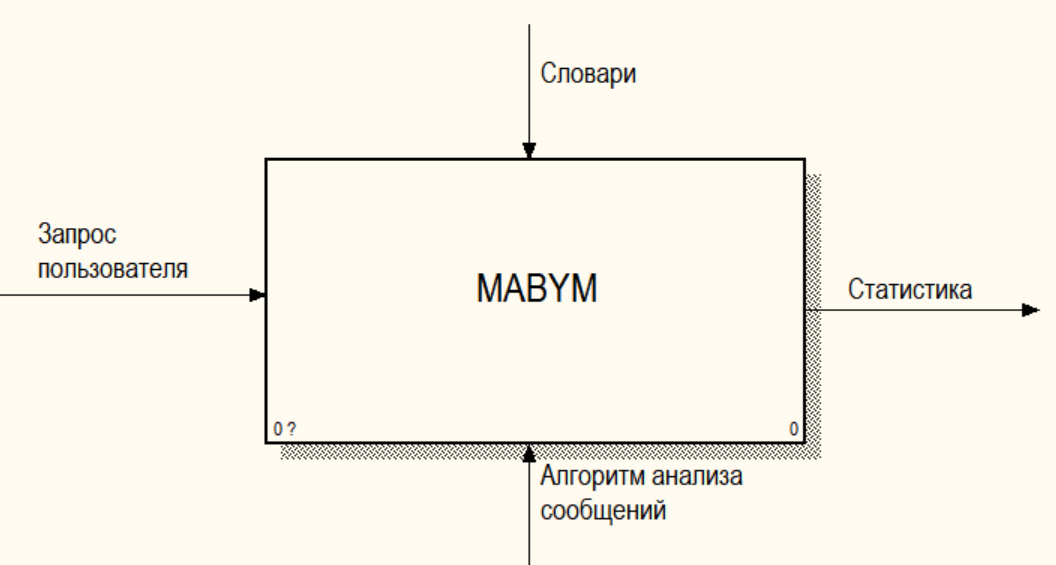


Рисунок 3. Контекстная диаграмма.

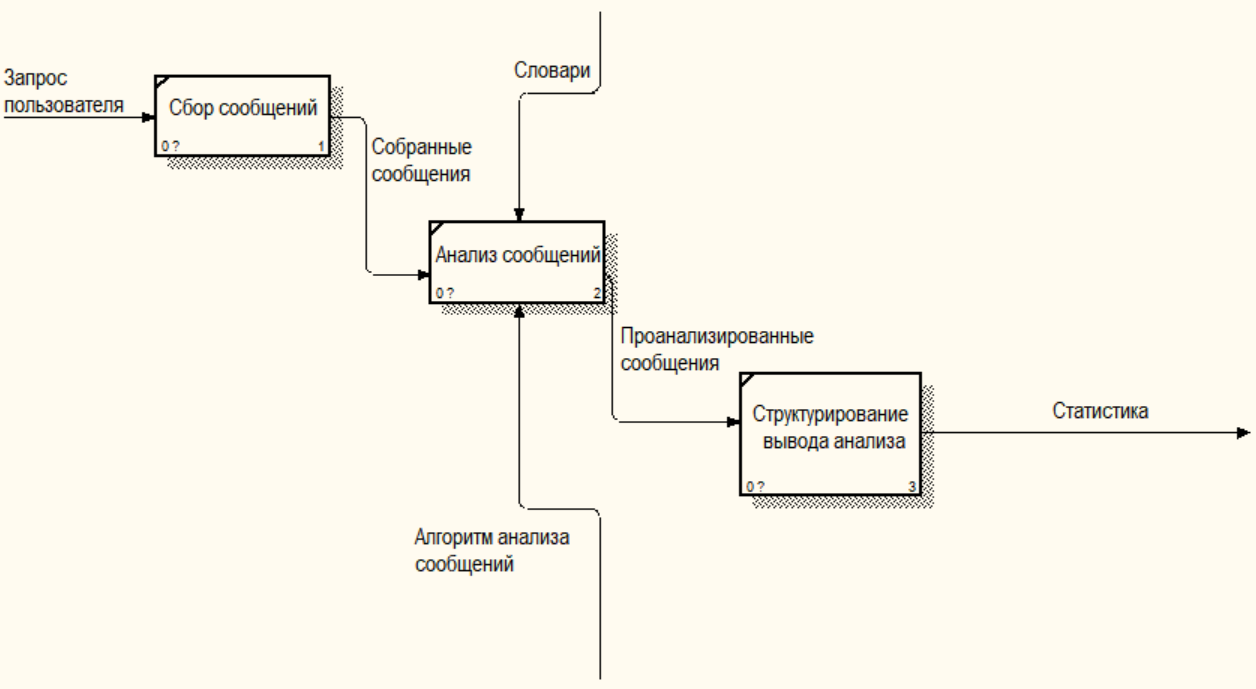


Рисунок 4. Декомпозиция контекстной диаграммы.