

## Оглавление

<b>РЕФЕРАТ</b>	<b>4</b>
<b>ВВЕДЕНИЕ</b>	<b>6</b>
<b>1 Аналитический раздел</b>	<b>7</b>
1.1 Интегративный подход в описании суицидального поведения . . . . .	7
1.2 Предотвращение самоубийств . . . . .	13
1.3 Выделение и классификация признаков паттернов суицидального поведения человека . . . . .	18
1.4 Форматы хранения проявление поведения человека . . . . .	21
1.5 Классификация сообщений в сети Интернет . . . . .	22
1.6 Формализация задачи метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям . . . . .	24
<b>2 Конструкторский раздел</b>	<b>28</b>
2.1 Формат и метод сбора данных . . . . .	28
2.2 Средства реализации ботов в мессенджере Telegram . . . . .	29
2.3 Декомпозиция системы . . . . .	29
2.4 Методы машинного обучения . . . . .	34
2.5 Алгоритмы векторизации текстовых сообщений . . . . .	39
<b>3 Технологический раздел</b>	<b>42</b>
3.1 Выбор инструментов разработки . . . . .	42
3.2 Интерфейс средства сбора данных . . . . .	43
3.3 Интерфейс средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям . . . . .	45
3.4 Описание обрабатываемых данных . . . . .	47
<b>4 Исследовательский раздел</b>	<b>52</b>
4.1 Условия исследований . . . . .	52
4.2 Исследование применимости моделей машинного обучения в методе распознавания суицидальных паттернов поведения человека по текстовым сообщениям . . . . .	52
4.2.1 Градиентный бустинг . . . . .	52
4.2.2 Метод случайного леса . . . . .	57
4.2.3 Метод опорных векторов . . . . .	62
4.2.4 Метод К-ближайших соседей . . . . .	67
4.2.5 Логистическая регрессия . . . . .	73
4.2.6 Перцептрон . . . . .	79

<b>ЗАКЛЮЧЕНИЕ</b>	<b>87</b>
<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>90</b>

# РЕФЕРАТ

Расчетно-пояснительная записка к выпускной квалификационной работе “Метод распознавания паттернов суицидального поведения человека по текстовым сообщениям” содержит 94 с., 49 рис., 3 табл., 50 ист., 0 прил.

Ключевые слова: суицидология, суицидальность, анализ текста, поведение человека, машинное обучение, классификация.

Объектом разработки является метод распознавания паттернов суицидального поведения человека.

Цель работы: разработка и программная реализация метода распознавания паттернов суицидального поведения человека по текстовым сообщениям.

В аналитическом разделе рассматриваются термины предметной области, представлен интегративный теоретический подход в суицидологии и приведена информация об определении истинности суицидальных намерений. Описываются факторы повышенного суицидального риска и мероприятия по предотвращению самоубийств. Приводится статистика совершения самоубийств. С использованием классификации признаков паттернов суицидального поведения человека описываются форматы хранения проявления поведения человека. Представлена информация о задействованных в настоящее время алгоритмах в задачах классификации сообщений в сети Интернет. Приводится формализация задачи метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям.

В конструкторском разделе описывается метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям, а также формат и метод сбора задействованных в нем данных. Рассматриваются средства реализации автоматизированного средства сбора данных. Приводится диаграмма вариантов использования, декомпозиция задачи распознавания суицидального сообщения, а также диаграмма “сущность-связь” в нотации Чена. Определяется перечень задействованных методов машинного обучения и векторизации, а также приводятся их схемы работы.

В технологическом разделе определяются инструменты разработки средства сбора данных и средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям. Реализован метод рас-

познавания суицидальных паттернов поведения человека по текстовым сообщениям. Представлены интерфейсы разработанных средств. Приводится описание обрабатываемых данных, а также анализ тональности сообщений и облаков слов каждого класса.

В исследовательском разделе проведено сравнительное исследование задействованных в методе алгоритмов машинного обучения и определено, какой из рассмотренных методов позволяет достичь лучших метрик точности с использованием матрицы ошибок и графиков оценок классификаторов.

Разработанный метод применим в сфере суицидологии. Автоматизированные средства поиска суицидальных сообщений, а также обычные анализаторы сообщений могут задействовать данный метод для обнаружения индивидов в суицидоопасных состояниях, а также для наблюдения людей, пребывающих в периоде пресуицида и постсуицида.

# **ВВЕДЕНИЕ**

По данным всемирной организации здравоохранения более 700 тысяч людей погибло от суицида в 2019 году. Суицид – это четвертая по частоте причина смерти молодых людей в возрасте от 15 до 29. [1]

Одним из ключевых эффективных мероприятий, основанных на этих данных, является раннее обнаружение и поддержка людей, подверженных суицидальным наклонностям.

Цель работы – разработать и реализовать метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям.

Для достижения поставленной цели потребуется:

- провести анализ действий и характеристик, позволяющих распознать паттерны суицидального поведения;
- классифицировать признаки паттернов суицидального поведения человека;
- описать способ получения набора данных суицидального поведения и разработать средство их сбора;
- разработать метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям;
- реализовать разработанный метод.

# 1. Аналитический раздел

В данном разделе рассматриваются термины предметной области, представлен интегративный теоретический подход в суицидологии и приведена информация об определении истинности суицидальных намерений. Описываются факторы повышенного суицидального риска и мероприятия по предотвращению самоубийств. Приводится статистика совершения самоубийств. С использованием классификации признаков паттернов суицидального поведения человека описываются форматы хранения проявления поведения человека. Представлена информация о задействованных в настоящее время алгоритмах в задачах классификации сообщений в сети Интернет. Приводится формализация задачи метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям.

## 1.1 Интегративный подход в описании суицидального поведения

Интегративный подход в психологии учитывает прежде всего способность человеческой психики меняться и совершенствоваться. В психотерапии интегративный подход направлен на процесс самораскрытия, саморазвития и самопомощь. [2]

Самоубийство – это умышленное лишения себя жизни [3]. На каждое самоубийство приходится гораздо больше несостоявшихся суицидентов, каждый несостоявшийся факт самоубийства является важным фактором риска самоубийства среди населения в целом [4].

Суицидной считается любая внутренняя либо внешняя активность, движимая стремлением человека лишить себя жизни. Внутренняя форма проявления суицидального поведения включает в себя пассивные суицидальные мысли, замыслы и намерения, а также соответствующий эмоциональный фон – суицидальные переживания. Внешнее суицидальное поведение проявляется в виде суицидальных практических суицидальных действий или высказываний, включая использование средств и способов, которые итоге могут привести к самоустранению или спровоцировать его попытку. [5]

**Суицидальным поведением** с точки зрения интегративного подхода можно считать любые внутренние и внешние формы психических актов, направляемые представлениями о лишении себя жизни. Внутренние формы

суицидального поведения включают: суицидальные мысли, представления, переживания, суицидальные тенденции, включающие замыслы и намерения. [6]

**Внешние проявления суицидального поведения** заключаются в попытках суицида и непосредственно завершённые суициды. В данном контексте под попыткой суицида понимается целенаправленное оперирование средствами лишения себя жизни, которое не закончилось смертью. [6]

Суицидальному поведению предшествуют **антивитальные переживания**, то есть размышления об отсутствии ценности жизни, где ещё нет четких представлений о собственной смерти, но имеется отрицание жизни [7]. Размышления об отсутствии ценности жизни выражаются в формулировках [6]: “не живу, а существую”, “бесцельное существование”.

Пассивные суицидальные мысли характеризуются представлениями и фантазиями на тему собственной смерти, но не в коем случае не лишения себя жизни как самопроизвольной активности. Зачастую подобное явление выражается в словах: “вот бы попасть под автобус”, “вот бы заснуть и не проснуться”. Суицидальные замыслы являются активной формой проявления суицидальности, то есть тенденция к самоубийству, причем она нарастает параллельно степени готовности плана ее реализации. В замыслах продумываются менее или более болезненные и медленные или быстрые способы лишения себя жизни, время и место действия. Суицидальные намерения предполагают присоединение к замыслу решения и волевого компонента, побуждающего к непосредственному переходу. **Пресуицид** – это период от возникновения суицидальных мыслей до попыток их реализации. При остром пресуициде период может исчисляться минутами, месяцами – при хроническом течении. [6]

В соответствии с целью суицидальное поведение подразделяют на истинное и парасуицидальное. **Истинные самоубийства**, покушения и тенденции имеют целью лишение себя жизни с некоторой целью, например, прекращение страданий или избавление от галлюцинаторных преследователей. **Парасуицид** зачастую используется для эпатажа окружающих в случае антисоциальных личностей и для шантажа в случае истероидных личностей. В случаях, когда попытка самоубийства при парасуициде заканчивается совершённым суицидом в следствие обстоятельств, а не желания суицидента, его смерть относится к несчастному случаю. При этом парасуицид встречается

ся у аддиктов с их жаждой острых ощущений, таким образом они пытаются вывести себя из состояния бесчувственности и безрадостности, к подобным актам можно отнести: прижигание кожи или удушение, хождение по перилам моста, демонстрация окружающим готовности покончить с собой для установления эмоционального контакта. В рамках аддиктивного поведения встречается также групповое и массовое самоубийство, совершаемое членами сект по религиозным мотивам. [6]

Личностный смысл суицидального поведения включает в себя выражение протеста, желание мести, стремление к признанию, избегание определенной ситуации, самонаказание и отказ [6]. Протест проявляется в случае конфликта с неприязненным объектом, на которого направлен акт самоубийства. Месть – это форма протеста, выражающаяся в ущемлении интересов противника. Такие виды поведения подразумевают наличие высокой самооценки и активную позицию личности со сменой направления агрессии извне внутрь. Суицидальное поведение может интерпретироваться как просьба о помощи в 90% случаев, и лишь в 10% случаев – как истинное желание покончить с собой [8].

Одной из самых уязвимых в плане суицидальной активности категорий являются подростки и молодежь. Более половины попыток суицида подростков являются демонстративными, но различить истинные и демонстративные суицидные попытки – одна из сложнейших задач. По клиническим данным, примерно каждый третий случай заверщенного суицида сопровождается неизвестной мотивацией. [5]

**Постсуицид** – период, следующий за попыткой суицида. Его подразделяют на 4 типа. [6]

*Критический постсуицид* характеризуется утратой актуальности конфликтом, отсутствием суицидальных тенденций, негативным отношением к совершенной попытке, чувством вины и стыда перед окружающими, страхом перед возможным смертельным исходом произведенной попытки и пониманием того, что это не могло исправить ситуации; помощь может быть ограничена рациональной психотерапией. [6]

*Манипулятивный постсуицид* характеризуется значительным улучшением конфликтной ситуации для пациента под влиянием его действий, отсутствием суицидальных тенденций, рентным отношением к совершенному,



легким чувством стыда и страхом перед возможным смертельным исходом. В данном случае суицидальное поведение закрепляется как способ воздействия на окружающих, кроме того закрепляется и демонстративно-шантажное поведение. В таких случаях требуется изменение ценностных ориентаций, разрушение шаблонов поведения, а также выработка негативного отношения к суициду. [6]

*Аналитический постсуицид* характеризуется неизменной актуальностью конфликта для суицидента, отсутствием суицидальных тенденций, негативным отношением к совершенной попытке. Используются новые способы разрешения конфликтной ситуации. При неэффективности или невозможности вынести конфликтную ситуацию возможна повторная попытка суицида с большим риском смертельного исхода. В данных случаях необходима помощь в ликвидации конфликта со стороны соответствующих служб. Если же эта ситуация обусловлена психопатологической продукцией (бред, галлюцинации), необходимо лечение и систематическое наблюдение психиатра. [6]

*Суицидально-фиксированный постсуицид* характеризуется актуальностью конфликта, сохранением или сокрытием суицидальных тенденций, положительным отношением к суициду. В данном случае требуется строгий надзор и лечение в условиях закрытого психиатрического стационара. [6]

Истинность суицидальных намерений зачастую определяются обстоятельствами попытки, субъективными сведениями, а также медицинскими критериями. [6]

В качестве *обстоятельств попытки* могут рассматриваться [6]:

- отсутствие в контакте с суицидентом окружающих лиц, а также малая вероятность появления кого-либо поблизости;
- время совершения действия от 6 до 12 часов дня;
- отсутствие суицидальных высказываний;
- отсутствие алкогольной интоксикации;
- приготовление к смерти: уборка личного пространства, смена личного белья;
- насильственные способы совершения суицида: попадание под транспортные средства, повешение, падение с высоты, использование огнестрельного оружия;

- принятие мер, препятствующих обнаружению или вмешательству: отключение личного средства связи, запирание двери на ключ и прочее. В качестве *субъективных сведений* о попытке выступают [6]:
- непосредственное желание умереть;
- длительность пресуицида (от суток и более);
- желание умереть и сожаление, что человек не смог довести дело до конца.

*Медицинские критерии* включают в себя: высокую вероятность смертельного исхода без медицинского вмешательства, а также отсутствие необходимости реанимационных мероприятий. [6]

Мотивы и поводы суицидальных поступков в порядке последовательного уменьшения их удельного веса [9]:

1. Лично-семейные конфликты:

- несправедливое отношение со стороны родственников и окружающих, включающее в себя оскорбления и унижения;
- супружеская измена или развод;
- половая несостоятельность;
- неудачная или безответная любовь;
- одиночество, социальная изоляция, изменение образа жизни;
- потеря близких или значимых знакомых, обнаружение у них болезней;
- препятствия к удовлетворению ситуационной актуальной потребности;

2. Состояние психического здоровья:

- реальные конфликты психически больных;
- патологические мотивировки;
- постановка психиатрического диагноза;

3. Состояние физического здоровья:

- соматические заболевания, физические страдания;
- хронические заболевания;
- уродства;

4. Конфликты, связанные с антисоциальным поведением суицидента:

- опасение судебной ответственности;
- боязнь наказания, позора, социального неприятия;

- самоосуждение;
- 5. Конфликты в профессиональной и учебной сфере:
  - падение престижа, несостоятельность;
  - несправедливые требования к выполнению работы;
  - несправедливая оценка исполненных обязанностей;
- 6. Материально-бытовые трудности, завышенные притязания;
- 7. Другие мотивы.

Среди факторов повышенного суицидального риска выделяют *экстраперсональные* и *внутриличностные*. К наиболее важным экстраперсональным факторам относят [10]:

- психозы и пограничные психические расстройства;
- суицидальные высказывания;
- повторные суицидальные действия и ранний постсуицидальный период до 3 месяцев;
- подростковый возраст;
- экстремальные условия жизни (тюремное заключение, маргинальное сообщество, одиночество);
- утрата престижа, особенно престижа в группе сверстников;
- конфликтная ситуация, психотравмы;
- алкоголизм и наркомания.

К внутриличностным факторам относят [10]:

- отсутствие установок, определяющих ценность жизни;
- неадекватность самооценки;
- неполноценность коммуникативной сферы;
- низкая стрессоустойчивость или толерантность к эмоциональным нагрузкам.

В основу рассматриваемой теории положена концепция самоубийства как следствия социально-психологической дезадаптации личности в условиях неразрешенного микросоциального конфликта. Процесс дезадаптации разделяется на фазы: *предиспозиционную* и *суицидальную*. Катализатором процесса перехода первой фазы во вторую является конфликт, который занимает центральное место в структуре суицидального акта, при этом суицид – это выход личности из конфликта путем самоустранения. Кроме того, суицид может представлять собой результат психического расстройства или

являться поведенческим актом практически здоровой личности. [6]

## **1.2 Предотвращение самоубийств**

Решение кризисных ситуаций без микросоциальной и эмоциональной поддержки, а также коррекции суицидогенных установок, практически невозможно. В ряде случаев кризисным пациентам необходим тренинг недостающих им навыков адаптации к сложившейся жизненной ситуации. Дифференцированный подход к суицидентам требует клинической квалификации формы суицидоопасной реакции. Таким образом, кризисные проблемы затрагивают ряд областей: психологию, педагогику и психиатрию. Поэтому ведущим методом помощи суицидентам должна быть психотерапия, однако традиционные методы психотерапии мало пригодны для решения кризисных проблем. [6]

Федеральный научно-методический центр суицидологии сыграл большую роль в организации в СССР разветвленной службы профилактики самоубийств. В соответствии с разработанной типологией суицида были созданы специализированные медико-психологические учреждения открытого типа, расположенные вне психиатрических больниц и ориентированные на различные диагностические группы. Также были разработаны организационно-методические основы функционирования таких подразделений суицидологической службы: телефон доверия, кабинет суицидолога, кабинет социально-психологической помощи районных поликлиник и кризисный стационар в многопрофильной городской больнице. Кроме того, была разработана и внедрена в суицидологическую практику дифференцированная программа психотерапии суицидентов, ориентированная на различные типы суицидоопасных состояний. Однако, к сожалению, в ходе социально-экономического кризиса 1990 годов, отечественная система превенции суицидов утратила свои прежние позиции. Во многих местах помощь суицидентам сводится к психиатрической госпитализации, что негуманно и неэффективно как с точки зрения затрачиваемых средств, так и терапевтически. Часто такая помощь заключается лишь в неотложных соматических мероприятиях, проводимых в отделениях скорой помощи, после выписки из которых суициденты часто совершают повторные попытки самоубийства. [6]

Одной из целей кризисной терапии является стремление избежать гос-

питализации, которая усложняет задачу адаптации пациента к его жизненному окружению. Поэтому пациентов без выраженного суицидального риска направляют в больницу лишь при наличии личной мотивации к стационарному лечению. Компромиссным вариантом является лечение в условиях дневного или ночного стационара.

Пресуицидальный синдром имеет следующие признаки [6]:

- резкое сужение интеллектуального фона, ограничение мыслительных процессов, заикленность мышления, ослабление способности рассматривать жизнеспособные варианты;
- сужение восприятия и уход в себя, чувство одиночества, бессмысленности и безвыходности;
- сильное смятение, ощущение крушения;
- бессильная агрессия и упреки в адрес других, сообщение о намерении покончить с собой;
- повышенная неприязнь к себе, самоотречение, ненависть, стыд, чувство вины;
- идея прекращения собственного существования, озарение, что существует способ положить конец собственным страданиям;
- фантазии о том, что будет после смерти с окружающими людьми.

Для оценки суицидального риска могут использоваться различные факторы, включающие в себя непосредственно суицидальную тематику, специфические симптомы и синдромы, а также влияние окружения. Для определения потребности в госпитализации или иных форм помощи пациентам разработаны и используются: шкала суицидального риска, карта риска суицидальности, шкала суицидального риска у больных алкоголизмом, опросник для больных с суицидальными тенденциями. Данные средства включают в себя анализ ситуации пациента и его личности. [6]

В шкале суицидального риска рассматриваются следующие факторы: возраст, пол, симптомы, стресс, суицидальное поведение в прошлом и настоящие планы, коммуникационные аспекты и возможности со стороны значимых других. При этом каждый из вариантов ответов имеет собственные баллы, по количеству которых и оценивается суицидальный риск. [6]

Карта для определения степени суицидального риска также имеет набор параметров, которые оцениваются некоторым количеством баллов, при-

чем наиболее значимые – баллами 2 и 3. [6]

В шкале суицидального риска больных алкоголизмом степень риска оценивается также алгебраической суммой баллов, соответствующих обнаруженным диагностическим признакам, причем низкий риск определяется как оценка ниже 400 баллов, средний – от 400 до 800 баллов, а высокий – от 800 до 1200 баллов, причем высокий риск служит показанием к госпитализации. [6]

Опросник для больных с суицидальными тенденциями основан на диагностике четырех ведущих сфер жизнедеятельности человека, включающих в себя тело и ощущения, профессию и деятельность, контакты и фантазии и будущее. [6]

Тактика кризисной терапии предполагает исследование значения стрессора для пациента, обеспечение необходимой социальной и микросоциальной поддержки, проявление сочувствия, побуждение к поиску альтернативных путей разрешения возникшей ситуации. При преобладании тревоги возможно применение гипноза. Может понадобиться вмешательство в виде внушения, переубеждения, изменения окружающей среды или госпитализации. Рекомендуются объяснить суициденту механизмы его саморазрушительного поведения и противопоставить его сознательные намерения и подсознательную мотивацию. Кроме того, может потребоваться выяснение социальных причин расстройства, а затем целостный анализ всех аспектов личности, после чего возможна смена ценностных ориентиров. Работа завершается замещением суицидальных установок “полезными”. [6]

Принято считать, что кризисный пациент утрачивает смысл своей жизни, однако вместо поиска нового смысла отказывается от жизни вообще. Основной целью кризисной терапии является использование пластичности коннотативной сферы суицидента, то есть его готовности отказаться от старой системы мышления и ценностей и выработать новые. Стратегия кризисной терапии заключается в создании психологических условий для личностного роста пациента. [6]

Поскольку кризис длится не более шести недель и локализуется в социально-психологической сфере, кризисная терапия должна быть краткосрочной. Кроме того, от нее требуется обеспечить индивиду практическую помощь в изменении позиции в кризисной ситуации.

Социально-психологическая направленность кризисной терапии отличается от других видов краткосрочной психотерапии и от методов межличностно-ориентированной психотерапии, которая используется для предупреждения рецидива кризиса у лиц с хроническим суицидальным риском. [6]

К целям кризисной терапии относят [11]:

- снятие симптомов;
- восстановление докризисного уровня функционирования;
- выявление внутренних ресурсов клиента, его семьи и иных форм помощи извне для преодоления кризисной ситуации;
- осознание событий, приводящих в состояние дисбаланса;
- осознание связи между стрессом и прежними жизненными переживаниями и проблемами;
- освоение новых моделей восприятия, мыслей чувств, развитие новых адаптивных реакций и стратегий совладания со стрессом, которые могут быть полезны не только в период данного кризиса, но и в будущем.

Главным показанием к применению кризисной терапии относят суицидоопасные состояния, обусловленные кризисной ситуацией, проявляющимися в сфере аффективных реакций субклинического и клинического уровня и развивающиеся у практически здоровых лиц, больных с пограничными нервно-психическими расстройствами и сохраненных душевнобольных вне связи с эндогенными механизмами заболевания. Кризисная терапия может осуществляться в индивидуальной, семейно и групповой формах, причем индивидуальная помощь является основой для других форм кризисной терапии и включает ряд этапов, которые в суицидологической практике могут частично перекрываться. [6]

Определен перечень действий, которые запрещены консультанту при проведении терапии [12]:

- впадать в замешательство или выглядеть шокированным;
- спорить или отговаривать пациента от суицида;
- преуменьшать боль пациента;
- пытаться улучшить и исправить состояние клиента, важно показать, что специалист понимает его боль;
- предлагать простых ответов на сложные вопросы, о проблемах человека надо говорить серьезно, открыто и откровенно, требуется оценивать

значимость этих проблем с точки зрения пациента;

- обесценивать опыт клиента путем упоминания чужих проблем;
- обещать держать план суицида в секрете.

Российская ассоциация телефонов экстренной психологической помощи является ассоциированным членом Международной федерации служб неотложной телефонной помощи IFOTES (International Federation of Telephone Emergency Services). Международные нормы федерации определяют цели, принципы и методы работы Телефонов Доверия. [6]

Службы неотложной телефонной помощи прилагают усилия, чтобы страдающий, отчаявшийся или думающий о самоубийстве человек имел возможность установить немедленный контакт с другим человеком, который готов его выслушать как друга и имеющим навыки оказания помощи в ходе беседы при уважении полной свободы абонента. Эта помощь распространяется не только на первый телефонный контакт, но продолжается в течение всего психологического кризиса, пока человеку требуется совет и поддержка. По желанию абонента служба может связать его с другим специалистом, компетентным в решении именно его проблемы. Любая помощь, оказанная службами телефонной неотложной помощи, имеет целью поддержать желание преодоления психологического кризиса. [6]

Для достижения положительных результатов абонент должен быть уверен в полной конфиденциальности, поэтому никакая информация, полученная от него, не может быть вынесена за пределы службы без его специального разрешения. Кроме того, ни абоненты, ни работники служб не должны подвергаться религиозному, политическому и иному давлению. Требуется, чтобы работники принимались в службы только после тщательного отбора и обучения. Кандидат в обязательном порядке должен обладать эмпатией. На абонента не возлагается никаких финансовых или иных обязательств при обращении в службу. [6]

Штат службы может состоять как из оплачиваемых работников, так и из волонтеров. Квалификация работников повышается путем постоянно продолжающегося обучения. Каждой службе должны быть доступны профессиональные консультанты разного профиля, которые могут быть как штатными сотрудниками службы, так и не состоять в ней вовсе. [6]

Психологическая помощь по телефону обладает рядом преимуществ,



которые, по сравнению с традиционной очной терапией, весьма значимы для кризисных пациентов [6]:

- пространственные – телефонная связь позволяет оказывать психотерапевтическое воздействие на абонента, находящегося на любом расстоянии от специалиста;
- временные – абонент может в любое удобное для него время незамедлительно соединиться с психотерапевтом, что особенно важно для лиц с низкой выносливостью к психологическому стрессу;
- анонимность – даже при обращении к врачу без предъявления документов анонимность обратившегося не является полной, ведь врач видит пациента и в дальнейшем может его узнать;
- возможность прервать контакт – телефонный абонент может прервать беседу с специалистом в любой момент, подобная возможность имеет большое значение для лиц, нуждающихся в психологической безопасности;
- эффект ограниченной коммуникации – исключительно акустический характер коммуникации способствует вербализации, тем самым и лучшему осознанию переживаемой ситуации, помогая пациенту интеллектуально овладеть травмирующей ситуацией.

Целью телефонной терапии является предотвращение дальнейшего развития острых кризисных состояний, помочь разрешить психотравмирующую ситуацию и тем самым предотвратить возможное покушение на самоубийство, что отвечает основным задачам данной терапии: помощь в овладении и преодолении актуальной психотравмирующей ситуации и коррекция неадаптивных личностных установок, обуславливающих развитие кризисных состояний и суицидальных тенденций.

### **1.3 Выделение и классификация признаков паттернов суицидального поведения человека**

Каждый год в мире совершается 703 000 самоубийств и еще больше попыток самоубийств. Данные акты происходят не только в странах с высоким уровнем дохода, но и являются глобальным явлением во всех регионах мира. [4]

В период с 2010 по 2021 год уровень самоубийств увеличился примерно на 36%. В 2021 году самоубийства стали причиной 48 183 смертей, также

12.3 миллиона взрослых американцев серьезно думали о самоубийстве, 3.5 миллиона планировали попытку самоубийства, а 1.7 миллионов человек пытались покончить жизнь самоубийством. [13]

На каждую смерть от самоубийства в 2021 году приходилось около [13]:

- 3 госпитализации по причине членовредительства;
- 8 обращений в отделение неотложной помощи в связи с самоубийством;
- 38 попыток самоубийства, о которых сообщили сами люди за последний год;
- 265 человек, которые всерьез задумывались о самоубийстве за последний год.

Также стоит отметить, что уровень самоубийств среди мужчин в 2021 году был примерно в 4 раза выше, чем среди женщин. Таким образом 50% населения является 80% самоубийц. Самый высокий же уровень самоубийств по возрасту наблюдается у людей старше 85 лет, таким образом, биологические признаки позволяют сужать область поиска в группах риска.

Статистика количества суицидов в определенные месяцы года представлена на рисунке 1.1 [13].

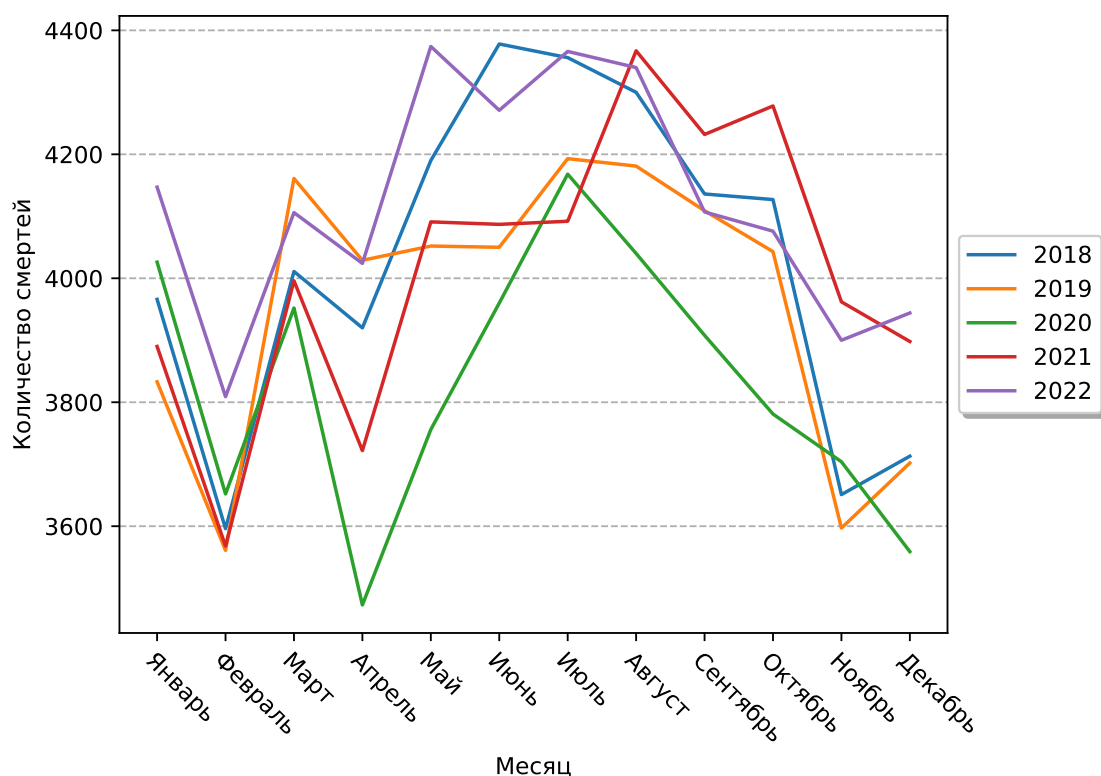


Рисунок 1.1 – График количества смертей от суицида в определенные месяцы за 2018–2022 год.

Признаки паттернов суицидального поведения человека по своей форме проявления могут быть *явными* и *неявными*. К явным признакам суицидального поведения относятся просьбы о помощи, в то время как к неявным – намеки на собственное состояние, депрессивные высказывания, изоляция от общества. Каждая форма проявления, как просьба о помощи, может выражаться:

- аудиально – с использованием речевого аппарата, интонаций, музыкальных композиций;
- в текстовом виде – записка или сообщение;
- визуально – с использованием собственного имиджа, жестов, паттернов движений;
- социально – через изоляцию или конфликты с обществом;
- физиологически – затяжная депрессия, приводящая к суицидальным наклонностям, сказывается на состоянии организма человека.

В сети интернет наиболее доступными являются явные и неявные признаки, выражающиеся в текстовом виде, причем в качестве анализируемого материала могут использоваться как личные сообщения, так и предоставля-

емые широкому кругу лиц тексты с описанием жизненных ситуаций, собственного состояния или личных переживаний.

#### **1.4 Форматы хранения проявление поведения человека**

Для автоматизации анализа выражений форм проявления суицидального поведения человека и выявления патологий требуется определить формат хранения данных.

Результатом анализа *аудиальных признаков* может быть наличие в словах человека депрессивных высказываний или интонаций пользователя в разговоре, умеющем суицидальный подтекст. Анализ аудио-сообщений подразумевает их расшифровку и сопоставление определенных высказываний с проявляемыми эмоциями. Далее под “эмоциональной картой аудио-дорожки” имеется в виду структура, в которой для определенных аудио-фрагментов сопоставлена одна из рассматриваемых эмоций пользователя. Данная структура может быть представлена как временные промежутки аудио-сообщения или как указатель на части расшифровки.

Результатом анализа *текстовых сообщений* может быть наличие в сообщении каких-либо суицидальных признаков: употребление выражений, подразумевающих суицидальную тематику, слова с суицидальным подтекстом, прямые высказывания о желании совершить суицид.

Результатом анализа *пространственно-временных признаков* может быть выявление подверженности жителей определенных регионов депрессии в связи с происходящими в них событиями или же соотнесение временных промежутков цикличным сезонам депрессии.

Результатом анализа *визуальных признаков* может быть определение проявления суицидального поведение, попытки суицида, либо его совершение, а также определение депрессии по мимике в течение разговора. Анализ визуальных признаков подразумевает обработку видеозаписей с действиями человека.

*Социальные признаки* могут выражаться как действия самого человека в социальной сфере или же информация, поступающая из социальной среды исследуемого индивидуума. Результатом анализа данных признаков может быть установление факта социальной изоляции, либо встревоженность окружения: родственников, друзей или знакомых. Анализ подразумевает мониторинг социальных связей человека или сбор данных о наблюдениях окру-

жающих индивида людей.

Анализ *физиологических признаков* может включать в себя либо диагностику депрессивного состояния, либо частотность присутствия исследуемого в стрессовых ситуациях, которые сопровождаются стадией истощения организма. *Общий адаптационный синдром* – это сочетание стереотипных реакций, возникающих в организме в ответ на действие стрессоров и обеспечивающих ему устойчивость не только к стрессорному агенту, но и по отношению к другим болезнетворным факторам [14]. Определены три стадии общего адаптационного синдрома: тревоги, устойчивости и истощения. При длительном воздействии стрессора адаптивные механизмы, участвующие в поддержании резистентности, истощаются, и наступает стадия истощения организма. Данная стадия не является обязательной и стрессовая ситуация может быть окончена до ее наступления. Протекание реакции заключается в активизации эндокринной системы и обеднении коры надпочечников. [14]

Анализ *биологических признаков* подразумевает под собой настройку чувствительности модели путем отнесения человека к известным группам риска. Согласно исследованиям в 2021 году констатировано 38 358 смертей в 2021 году и 39255 в 2022 году в результате суицида среди мужчин [13]. Среди женщин зафиксировано 9 825 смертей в 2021 году и 10 194 смертей в 2022 году [13]. Таким образом, по данной статистике мужчины совершают  $\approx 80\%$  самоубийств, что говорит о том, что пол имеет важную роль в определении суицидального поведения индивида. Также, согласно исследованиям чаще всего суицид совершают люди возрастной группы 25 -44 лет и 45 -64 лет, в 2021 году зафиксировано, что они составили совместно для этих групп  $\approx 65\%$  всех самоубийств в США, а в 2022 году  $\approx 66\%$  [13].

### **1.5 Классификация сообщений в сети Интернет**

В настоящее время в сети Интернет большое количество событий, явлений и процессов описывается в относительно коротких текстовых сообщениях до 3 000 символов. Их основными источниками являются сетевые новостные агентства. Накопленные информационными службами массивы таких сообщений активно используются при наблюдении за обстановкой, состояниями объектов, общественным мнением. Современное развитие сетевых СМИ, а также средств автоматического сбора текстовой информации

сети Интернет позволяет информационным службам получать тысячи сообщений в день. Так, только одно сетевое новостное агентство федерального уровня, как правило, выпускает несколько десятков сообщений в день, а роботизированные системы сбора сообщений позволяют обрабатывать сотни подобных источников. Информационные потоки большой интенсивности и размерности делают невозможным ознакомление эксперта с каждым сообщением и вычленением его семантики. Одним из решений данной проблемы является использование средств автоматической классификации текстовых сообщений, а необходимым условием эффективной работы подобным систем является применение алгоритма, обеспечивающего необходимое качество классификации. [15]

Популяризация и развитие всемирной паутины и расширение глобальных компьютерных сетей, в том числе пополнение информационных баз данных, повлекло за собой наращивание текстовых ресурсов. При интенсивном и постоянном возрастании объемов текстовой информации сложность поиска требуемых сведений среди огромного количества доступных текстов и сообщений значительно уменьшают ее ценность. Классификация сообщений необходима для решения следующих задач [16]:

- обнаружение спама;
- разделение сайтов по тематическим каталогам;
- распознавание эмоциональной окраски текстов;
- автоматизация процесса исследования естественного языка;
- генерация текста;
- автоматическая обработка текста на естественном языке.

Существующие подходы к решению задач анализа текстов и обработки естественного языка включают в себя [16]:

- эмпирические методы, основанные на вероятностных моделях машинного обучения в области обработки естественного языка [17];
- выявление психических заболеваний с помощью обработки естественного языка, выраженных в разнообразных текстовых данных, включая сообщения в социальных сетях [18];
- статистика достоверности электронных медицинских карт и восстановления показателей из плохо читаемых сообщений при помощи методов анализа естественного языка [19];

- взаимосвязь между эмоциями и настроениями клиентов, возникающими при взаимодействии с роботами в отелях с использованием методов анализа текста [20];
- исследование тональности финансовой информации на английском языке при помощи метода анализа больших данных и нейронных сетей [21];
- анализ настроений и эмоций для выявления стресса человека на основе сообщений и комментариев при помощи машинного обучения [22].

Задача определения тональности сообщений также может быть отнесена к задаче классификации сообщений, однако она не может решать более сложные проблемы, требующие изменения не только системы определения тональности. Зачастую при ее решении задействуют нейронные сети [16], либо различные методы машинного обучения [23]. В качестве методов машинного обучения зачастую рассматриваются [23]:

- изоляционный лес;
- локальный уровень выброса;
- логистическая регрессия;
- OneClassSVM;
- случайный лес;
- XGBoost [24].

Классификация сообщений необходима в организации социальных сетей для соблюдения порядка и своевременного обнаружения сообщений, разжигающих национальную рознь, призывающих к беспорядкам или оскорбляющих правительственную систему. Также известно о наличии в некоторых социальных сетях и искусственного интеллекта, предназначенного для обнаружения людей с высоким суицидальным риском [25], однако источники не представляются, а все наработки являются объектами коммерческой тайны.

### **1.6 Формализация задачи метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям**

Метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям включает в себя хранение и анализ сообщений пользователей. Для определения, является ли сообщение суицидальным, используется модель машинного обучения. В качестве обучающей выборки используется дополненный датасет размеченных сообщений из открытого доступа

[26].

На рисунке 1.2 представлена IDEF0 диаграмма нулевого уровня задачи определения наличия суицидальных паттернов в текстовом сообщении.

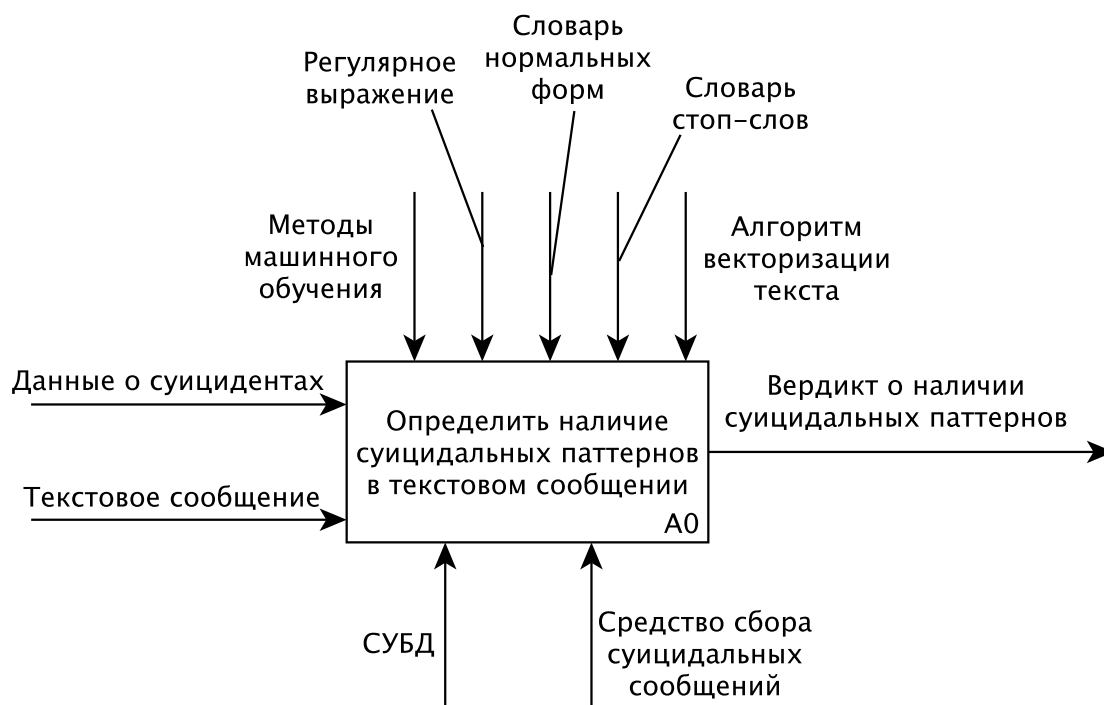


Рисунок 1.2 – IDEF0 диаграмма нулевого уровня.

Представлена информация о задействованных в настоящее время алгоритмах в задачах классификации сообщений в сети Интернет. Приводится формализация задачи метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям.

### Вывод

Были рассмотрены термины предметной области, включающие в себя понятия самоубийства и суицидального поведения. Представлен интегративный теоретический подход в суицидологии, описаны суицидальное поведение, внешние проявления суицидального поведения, пресуицид, парасуицид и постсуицид. Приведена информация об определении истинности суицидальных намерений, включившая в себя обстоятельства попытки и медицинские критерии.

Описаны факторы повышенного суицидального риска и мероприятия и методики предотвращения самоубийств, включающие в себя суицидологию



ческую диагностику, кризисную терапию и Телефон Доверия. Представлена статистика совершения самоубийств.

С использованием классификации признаков паттернов суицидального поведения человека описаны форматы хранения проявления поведения человека, представленные в таблице 1. Были выделены аудиальные, текстовые, пространственно-временные, визуальные, физиологические и биологические признаки. В рамках разрабатываемого метода анализируются лишь текстовые признаки.

Приведена информация о задействованных в настоящее время алгоритмах в задачах классификации сообщений в сети Интернет, которые включили в себя: изоляционный лес, локальный уровень выброса, логистическую регрессию, OneClassSVM, случайный лес и XGBoost.

Формализована задача метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям. Метод включил в себя хранение и анализ сообщений пользователей и определение суицидальных сообщений с использованием модели машинного обучения.

Таблица 1 – Форматы описания признаков и их методов обработки

Признаки	Данные	Методы обработки
Аудиальные	аудиофайл	распознавание речи, обработка и анализ текстовых сообщений, анализ характеристик голоса
	текстовая расшифровка речи	обработка и анализ текстовых сообщений
	эмоциональная карта, аудиофайл / текстовая расшифровка	сопоставление эмоциональной карты смысловой нагрузке речи
Текстовые	текстовое сообщение	обработка и анализ текстовых сообщений с использованием методов машинного обучения
	текстовое сообщение, эмоциональная карта	уточнение входных данных модели с использованием эмоциональной карты
Пространственно-временные	дата написания	соотнесения дат действий пользователя сезонности депрессии
	место дислокации автора, дата написания	соотнесение контекста происходящего в регионе пользователя его действиям
Визуальные	видеоряд действий пользователя	распознавание эмоций
	видеоряд действий пользователя, мониторинг контекста происходящего	анализ реакций индивидуума на внешние раздражители и жизненные ситуации
Физиологические	го	
	данные мониторинга уровня стресса	анализ состояния организма человека и его подверженности стрессам
	данные мониторинга уровня кортизола в крови	
Биологические	данные мониторинга состояния здоровья человека	
	пол пользователя	уточнение входных данных модели с использованием пола пользователя
	возраст пользователя	уточнение входных данных модели с использованием возрастной группы пользователя

## 2. Конструкторский раздел

В разделе описывается метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям, а также формат и метод сбора задействованных в нем данных. Рассматриваются средства реализации автоматизированного средства сбора данных. Приводится диаграмма вариантов использования, декомпозиция задачи распознавания суицидального сообщения, а также диаграмма “сущность-связь” в нотации Чена. Определяется перечень задействованных методов машинного обучения и векторизации, а также приводятся их схемы работы.

### 2.1 Формат и метод сбора данных

В качестве задействованных в анализе данных используются текстовые сообщения. Для сбора данных потребуется использовать автоматизированное средство сбора суицидальных сообщений в мессенджере Telegram. Интерфейс программного обеспечения позволяет направить в систему хранения два типа сообщений: суицидальные и на суицидальную тематику.

Средство сбора должно предоставлять пользователю следующий функционал:

1. Получение информации о проекте;
2. Получение информации об отличиях суицидальных сообщений и сообщений на суицидальную тематику;
3. Направление примеров суицидальных сообщений;
4. Направление примеров сообщений на суицидальную тематику, но не относящихся к суицидальным.

Согласно 152-ФЗ “О персональных данных”, “персональные данные – любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу (субъекту персональных данных)” [27]. Таким образом, к персональным данным можно отнести фамилию, имя и отчество, дату и место рождения, адрес проживания, семейное, социальное и имущественное положение, образование, профессию, доходы и другое. В связи с этим средство сбора информации не обрабатывает и не хранит никаких персональных данных о пользователях, направивших сообщения.

## 2.2 Средства реализации ботов в мессенджере Telegram

В качестве представленных к использованию в качестве средства реализации бота в мессенджере телеграм могут быть задействованы популярные библиотеки:

- Python Telegram Bot [28],
- Telebot [29],
- Node Telegram Bot API [30],
- Telegram Bot Kotlin [31].

Python Telegram Bot – это библиотека, предоставляющая асинхронный интерфейс на ЯП Python для Telegram Bot API, которая совместима с версиями Python3.8 [32] и выше. Данная библиотека предоставляет высокий уровень абстракции и позволяет использовать объектно-ориентированный подход. Помимо реализации API, она также содержит ряд классов высокого уровня, упрощающих разработку ботов. Проект поддерживает строенную с асинхронным вводом-выводом. [28]

Telebot – библиотека на Python, содержащая в себе асинхронную и синхронную реализацию Telegram Bot API. Данный проект предоставляет более гибкий и низкоуровневый доступ к API Telegram, чем упомянутый Python Telegram Bot. [29]

Node Telegram Bot API – это библиотека для создания Telegram-ботов с использованием языка JavaScript [33] и платформы NodeJS [34]. Несмотря на распространенное использование, используемая версия API Telegram является устаревшей. [30]

Telegram Bot Kotlin – это библиотека для создания Telegram-ботов на ЯП Kotlin [35]. Отличительной особенностью является возможность разработки ботов на платформе Java Virtual Machine [36], а также поддержка Kotlin Coroutines. [31]

## 2.3 Декомпозиция системы

На рисунке 2.1 представлена диаграмма вариантов использования системы.

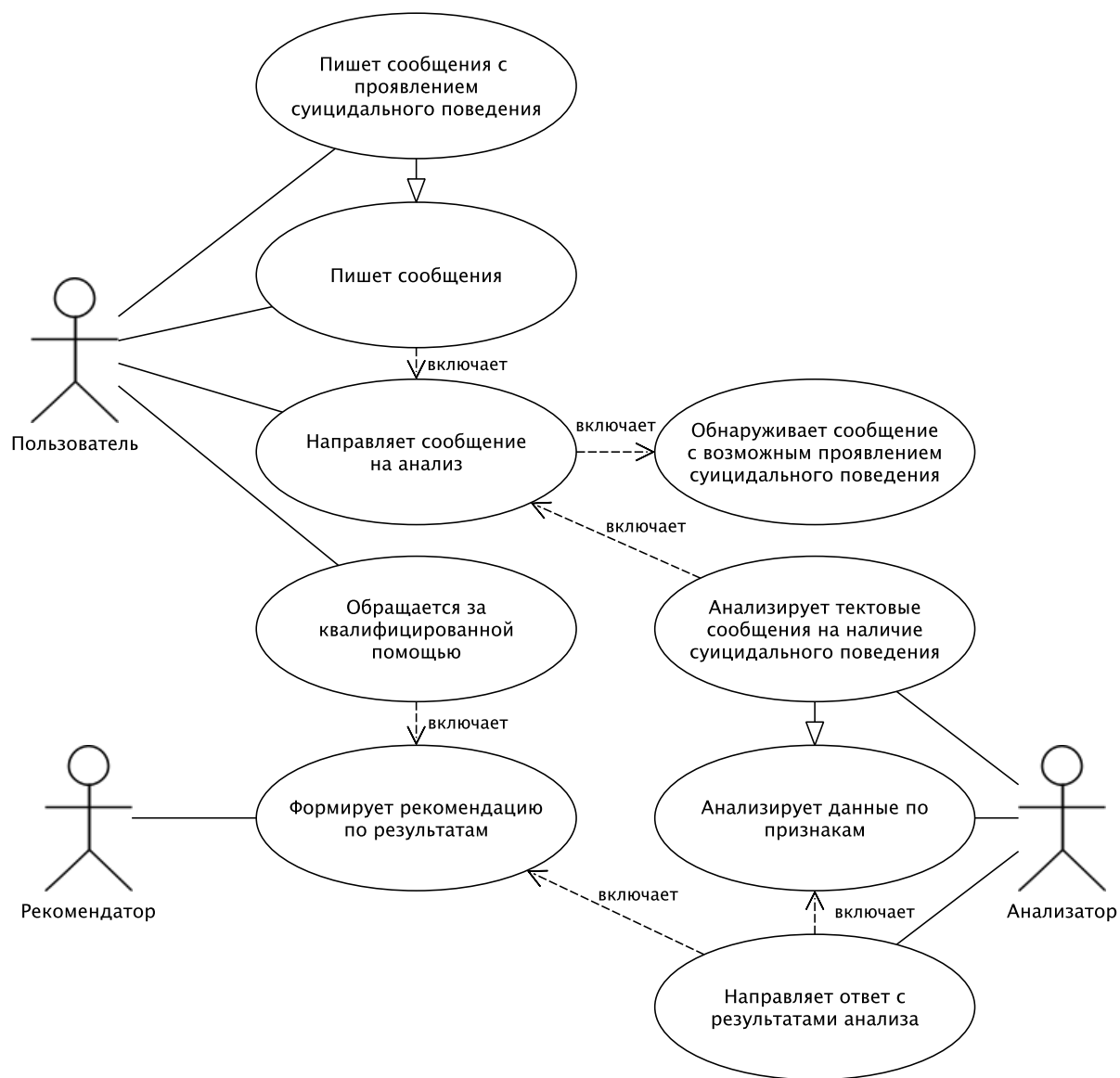


Рисунок 2.1 – Диаграмма вариантов использования системы.

На рисунке 2.2 представлена IDEF0 диаграмма первого уровня задачи определения наличия суицидальных паттернов в текстовом сообщении.

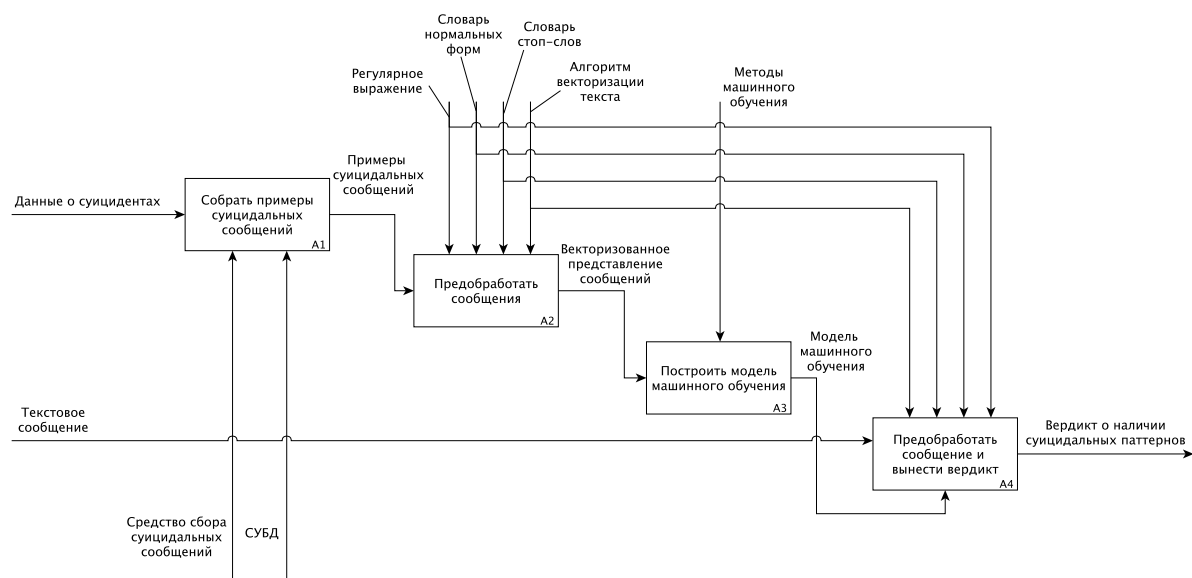


Рисунок 2.2 – IDEF0 диаграмма первого уровня.

Модуль A1 на рисунке 2.2 отвечает за сбор примеров суицидальных сообщений с использованием средства сбора суицидальных сообщений и СУБД. Средство сбора суицидальных сообщений в данном контексте подразумевает программное обеспечение, позволяющее пользователям вносить размеченные сообщения, то есть текст с указанным классом, к которому он относится. Затем эти данные будут задействованы в блоке предобработки сообщений.

Модуль A2 на рисунке 2.2 отвечает за предобработку данных, собранных на этапах работы модуля A1. Данный блок включает в себя следующие шаги: токенизацию, лемматизацию, удаление стоп-слов и векторизацию. В результате выполнения блока будет получено векторизованное представление поступивших в него примеров суицидальных сообщений, которое будет задействовано в дальнейшем в блоке построения модели машинного обучения. Декомпозиция рассматриваемого блока приведена на рисунке 2.3.



Рисунок 2.3 – IDEF0 диаграмма, декомпозиция блока A2.

Модуль A3 на рисунке 2.2 отвечает за построение модели машинного обучения, играющую роль классификатора сообщений. В качестве контроля блока выступают методы машинного обучения: градиентный бустинг, метод случайного леса, метод опорных векторов, метод К-ближайших соседей, логистическая регрессия и перцептрон. В результате выполнения будет получена модель машинного обучения, задействованная в блоке предобработки поступившего вне обучающей выборки сообщения и вынесения вердикта о его суицидальности.

Модуль A4 на рисунке 2.2 отвечает за предобработку поступившего в систему сообщения, а также его оценку на суицидальность. В качестве механизмов используются словарь нормальных форм, регулярное выражение и словарь стоп-слов для выполнения предобработки, а также модель машинного обучения для оценки сообщения.

На рисунке 2.4 представлена диаграмма “сущность-связь” в нотации Чена.

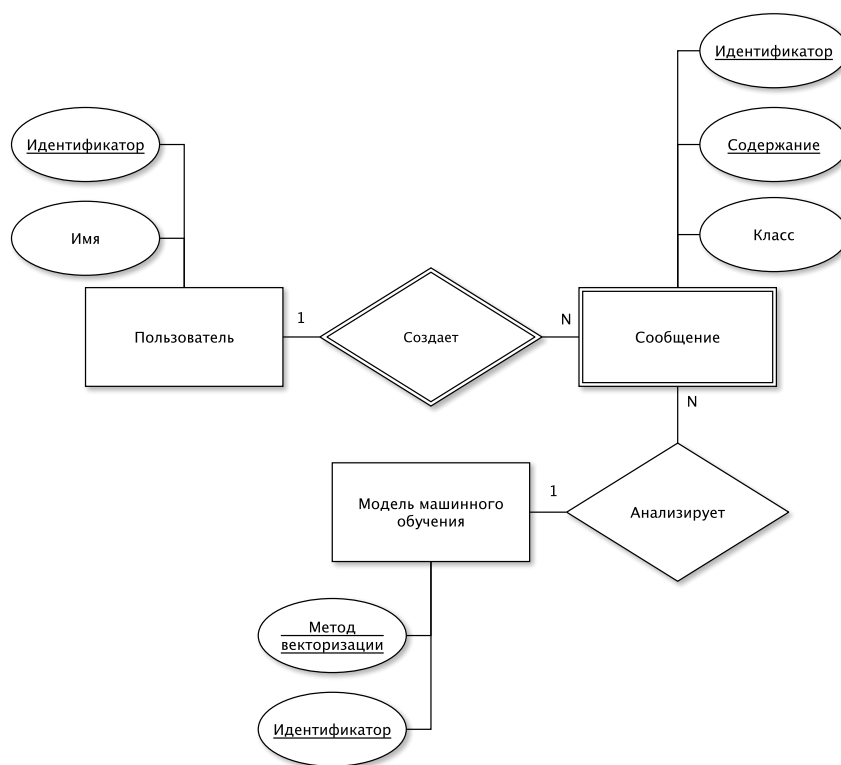


Рисунок 2.4 – Диаграмма “сущность-связь” в нотации Чена.

На рисунке 2.5 представлена схема программного обеспечения.

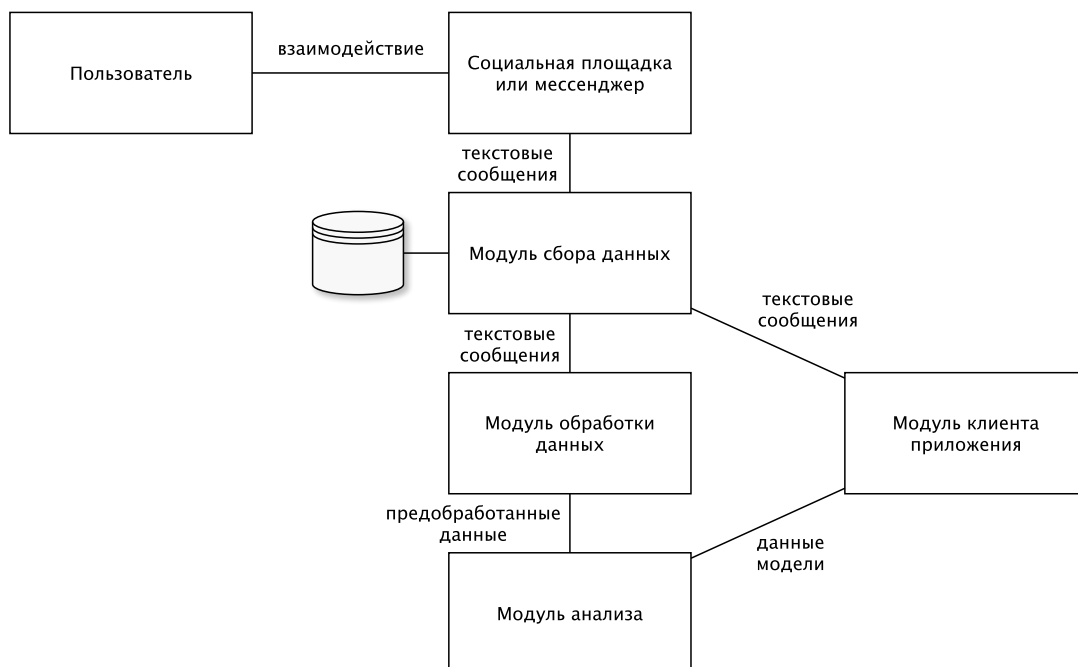


Рисунок 2.5 – Схема программного обеспечения.



## **2.4 Методы машинного обучения**

В качестве задействованных в методе алгоритмов рассматриваются: градиентный бустинг (схема алгоритма представлена на рисунке 2.6), метод случайного леса (схема алгоритма представлена на рисунке 2.7), метод опорных векторов (схема алгоритма представлена на рисунке 2.8), метод К-ближайших соседей (схема алгоритма представлена на рисунке 2.9), логистическая регрессия (схема алгоритма представлена на рисунке 2.10) и перцептрон (схема алгоритма представлена на рисунке 2.11). Данный список обусловлен потребностью в целях исследования охватить широкий спектр методов машинного обучения для дальнейшего изучения возможности и потребности в создании ансамблевых моделей для решения поставленной задачи.

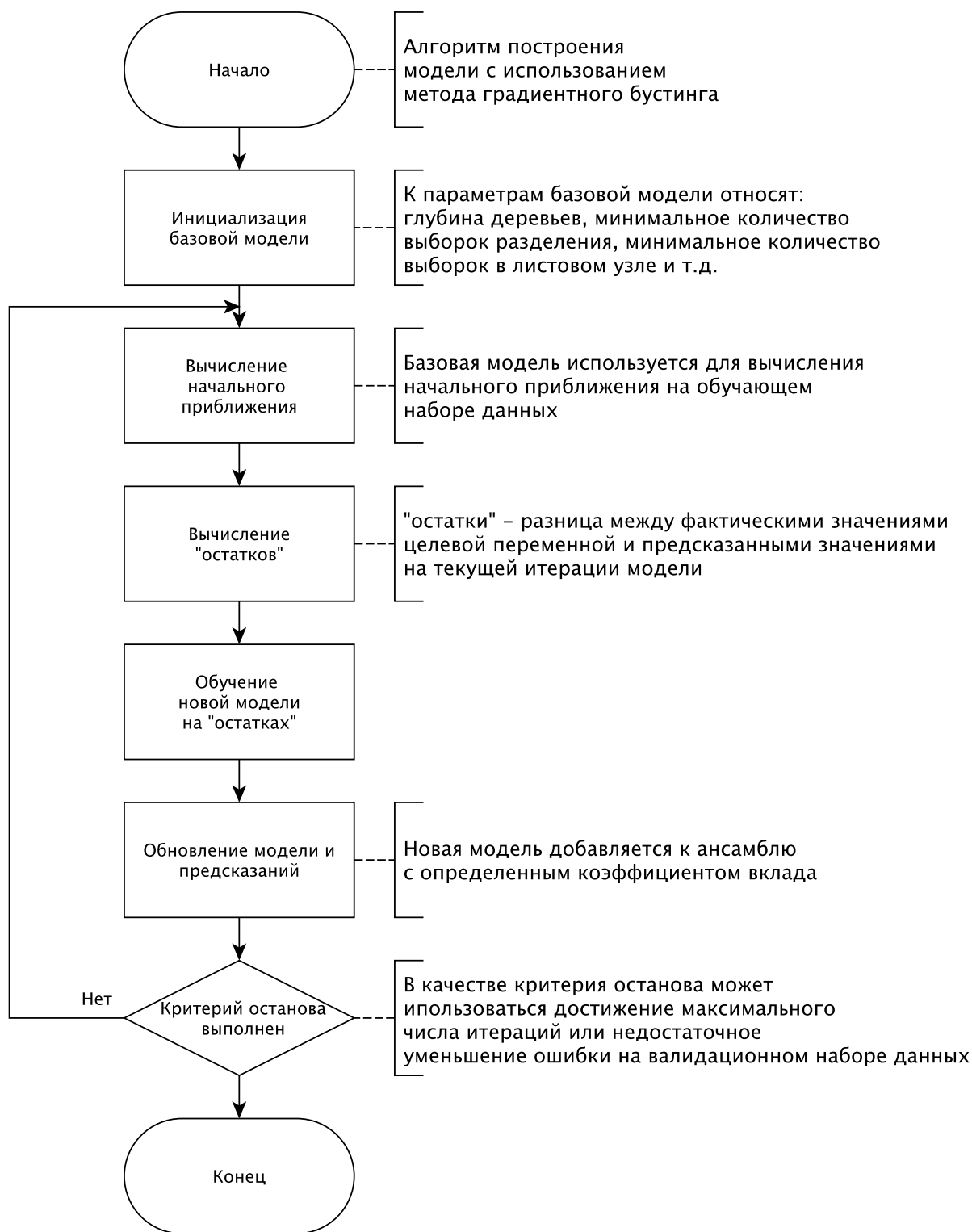


Рисунок 2.6 – Схема алгоритма работы градиентного бустинга.

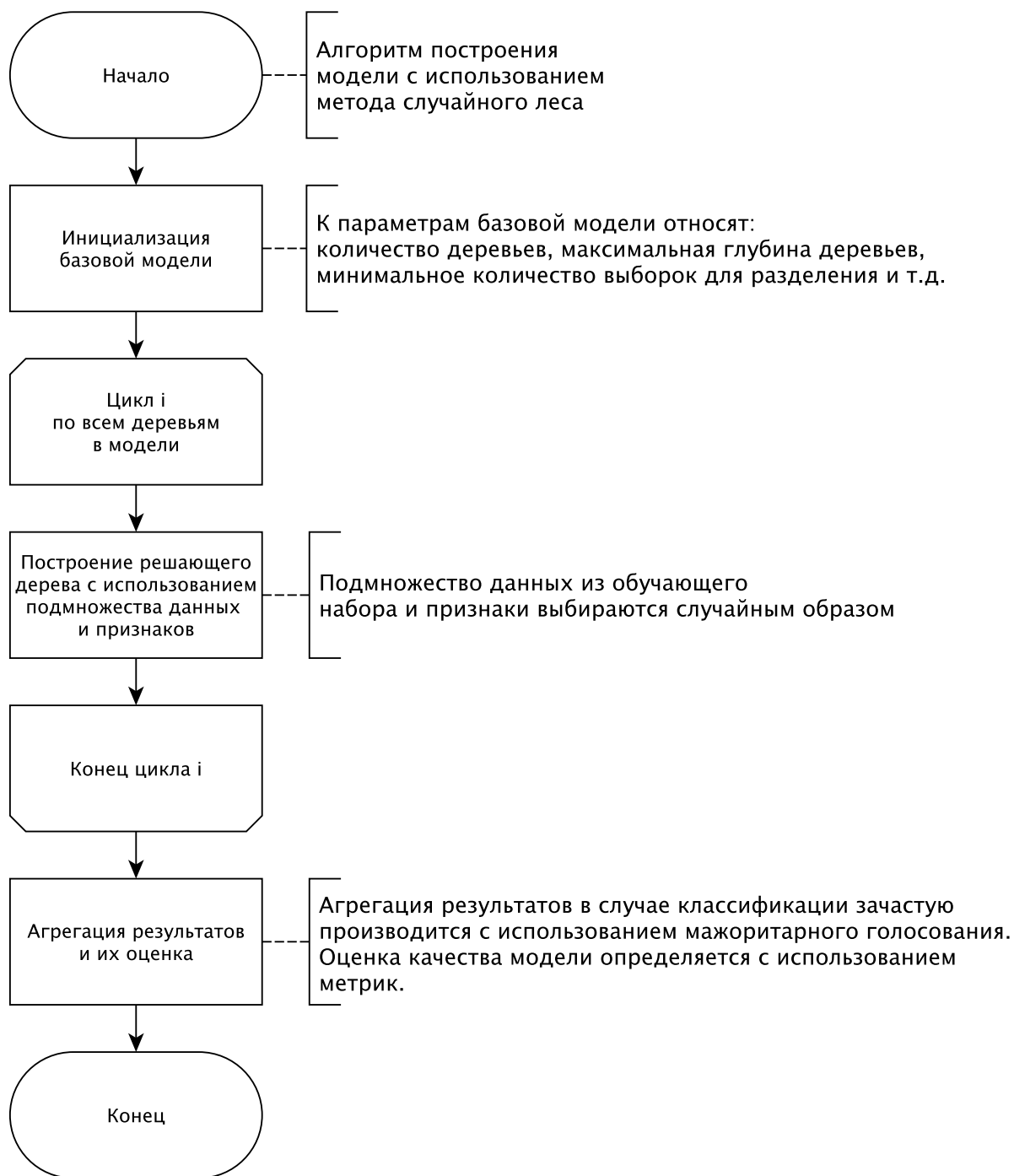


Рисунок 2.7 – Схема алгоритма работы метода случайного леса.

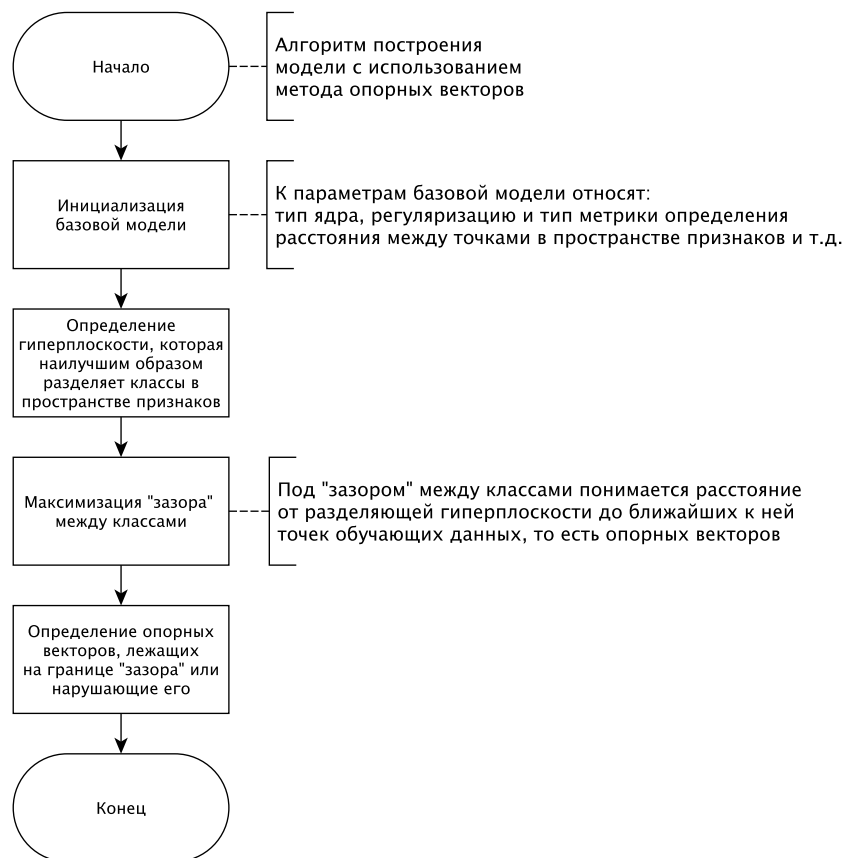


Рисунок 2.8 – Схема алгоритма работы метода опорных векторов.

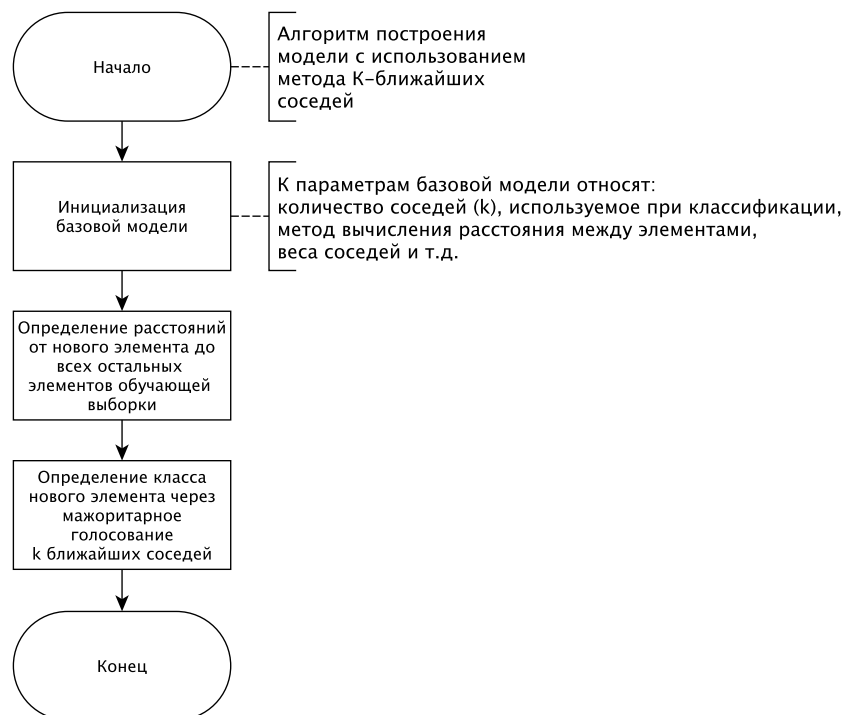


Рисунок 2.9 – Схема алгоритма работы метода K-ближайших соседей.

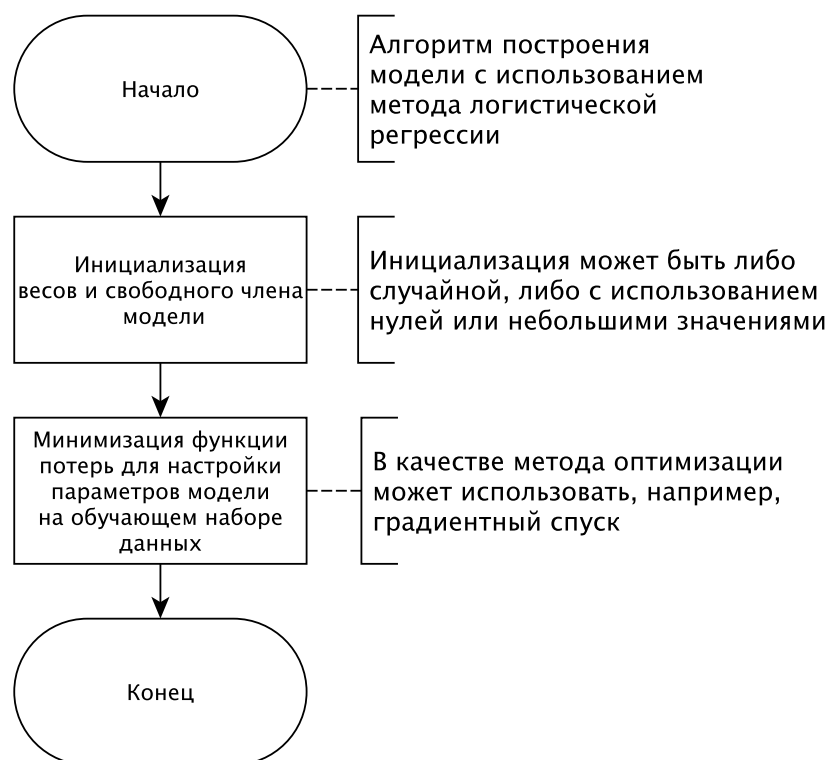


Рисунок 2.10 – Схема алгоритма логистической регрессии.

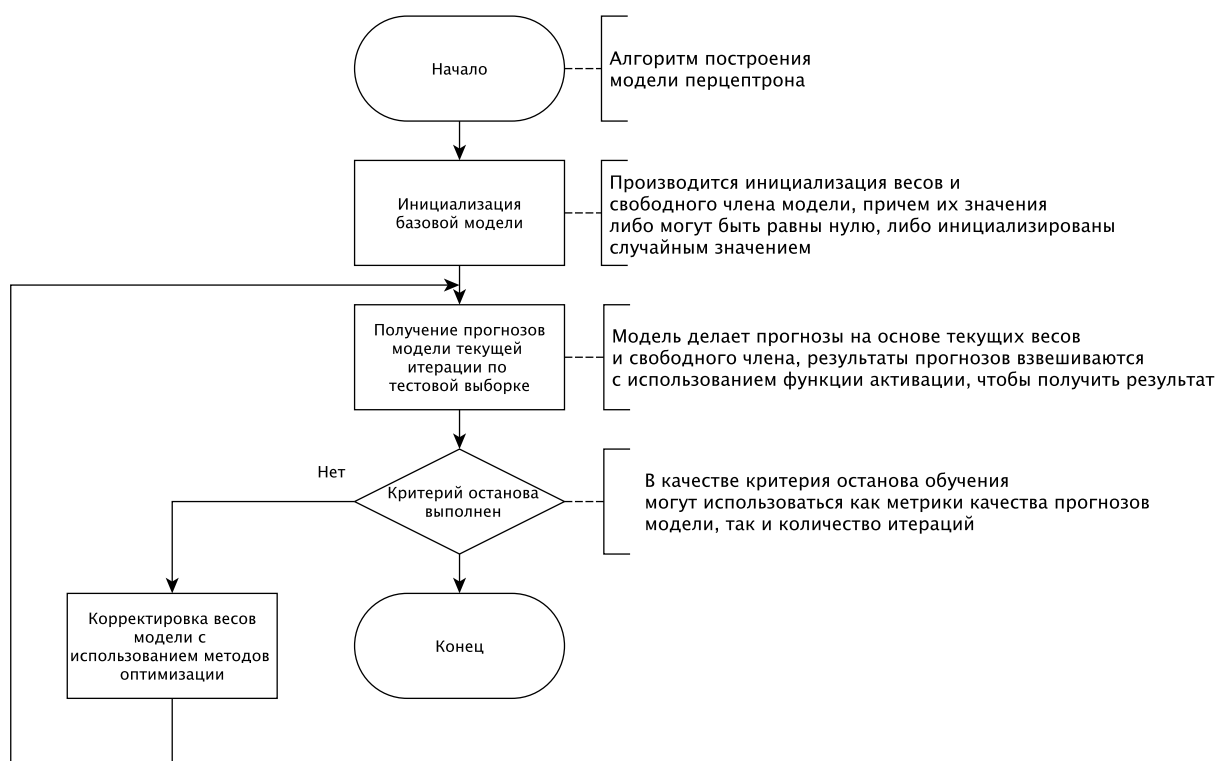


Рисунок 2.11 – Схема алгоритма работы метода с использованием перцептрона.

## 2.5 Алгоритмы векторизации текстовых сообщений

Классификация текстовых сообщений может производиться с использованием выделения ключевых слов, либо с использованием векторного представления данных. Первый подход является наиболее простым, но к его недостаткам относят необходимость участия экспертов в создании словарей для описания каждого классифицируемого класса сообщений, кроме того для получения приемлемых результатов требуется, чтобы классы сообщений мало пересекались по словарному набору [37]. Второй подход в классификации является более сложным, в нём могут использоваться различные модели векторизации для дальнейшей обработки текстовой информации. В качестве задействованной здесь может выступать модель текста “мешок слов” [38] с его расширением для частотных характеристик встречаемости слов в сообщениях TF-IDF [39]. Метод базируется на создании векторов сообщений с учетом весов встречаемости каждого слова, как в самом сообщении, так и во всех сообщениях выборки. Его использование при построении классификатора значительно повышает точность классификации в некотором наборе задач [37].

Кроме того, в качестве способа представления могут быть задействованы векторизации вида word embeddings. В качестве основного механизма они задействуют модель word2vec [40], которая представляет собой нейронную сеть, ставящую подаваемому на вход слову в соответствие выходной вектор заданной длины. Обучение нейронной сети производится таким образом, чтобы получить близкие в пространстве векторы для слов, встречающихся в одинаковых контекстах. Таким образом, слова близкие по значению, либо употребляющиеся совместно, будут иметь близкие векторы в пространстве. Существенным недостатком данного метода заключается в необходимости обучения нейронной сети на значительном корпусе текстов, что является ресурсозатратной задачей. Эта проблема нивелируется существованием предобученных моделей. Также к недостаткам относят невозможность представления слова, которого не было в обучающем корпусе. [37]

К векторизации word embeddings также относят и проект ELMo [41]. Данная модель обладает преимуществами word2vec и дополняет их возможностью формирования одного вектора на все сообщения, а также получения вектора для неизвестного слова путем его разложения на отдельные слоги и буквы. В основе метода лежит многослойная нейронная сеть, которая на первом

слое получает векторы для букв, а затем слои для работы со словами в составе всего сообщения. В качестве вектора используются веса последнего слоя сети. Для данного подхода также существуют предобученные модели для многих языков, что значительно упрощает его использование. [37]

BERT [42] – языковая модель, применяющаяся для анализа текста. Ее применение для построения классификатора требует обучения модели с кодированием отдельным слоем нейронной сети, что не дает возможность использовать данную технологию только для векторизации. В отличие от прежних классических языковых моделей, BERT обучает контекстно-зависимые представления, причем он учитывает двусторонний контекст, что помогает модели лучше понимать смысл многозначных слов. [37]

Для векторизации собранных текстовых сообщений будут задействованы алгоритм “мешок слов” и языковая модель BERT. Схемы работы выбранных алгоритмов представлены на рисунках 2.12 и 2.13 соответственно. Алгоритм TF-IDF не входит в рассмотрение в силу того, что он является модернизацией алгоритма “мешок слов”. Модели word2vec и ELMo также не попадают в рассмотрение в связи с тем, что данная работа первично имеет цель определить наиболее подходящий для решения задачи алгоритм машинного обучения.



Рисунок 2.12 – Схема работы алгоритма “мешок слов”.

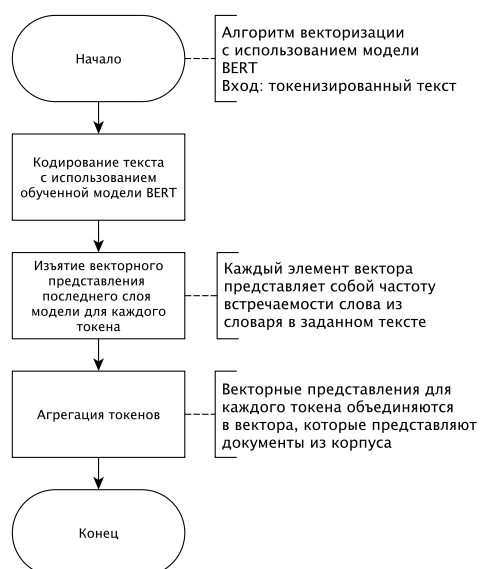


Рисунок 2.13 – Схема алгоритма векторизации с использованием модели BERT.

## Вывод

Был описан метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям, а также формат и метод сбора задействованных в нем данных. Определено, что в качестве средства сбора данных будет использоваться бот в мессенджере Telegram. Рассмотрены доступные средства реализации ботов в выбранном мессенджере.

Приведена диаграмма вариантов использования. Для системы было определено три действующих лица: пользователь, рекомендатор и анализатор. Приведена IDEF0 диаграмма, декомпозирована главная задача метода – распознавание суицидального сообщения. Диаграмма “сущность-связь” в нотации Чена позволила на абстрактном уровне описать систему распознавания.

Был определен перечень задействованных методов машинного обучения, который включил в себя: градиентный бустинг, метод случайного леса, метод опорных векторов, метод К-ближайших соседей, логистическая регрессия и перцептрон. В качестве методов векторизации выбраны: алгоритм “мешок слов” и языковая модель BERT. Для каждого выбранного алгоритма приведена схема его работы.



### 3. Технологический раздел

В данном разделе определяются инструменты разработки средства сбора данных и средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям. Представлены интерфейсы разработанных средств.

Приводится описание обрабатываемых данных, а также анализ тональности сообщений и облаков слов каждого класса.

#### 3.1 Выбор инструментов разработки

В качестве средства реализации средства сбора данных была выбрана библиотека Telebot, так как:

1. Функционал приложения не предусматривает сложных операций, в силу чего низкая производительность ЯП Python не скажется на скорости отклика системы;
2. ЯП Python позволит быстро разворачивать приложение на разнообразных операционных системах, поддерживающих интерпретатор Python;
3. Telebot предоставляет более тонкую настройку и контроль над запросами и ответами API Telegram.

Для организации хранения данных и моделей задействована реляционная СУБД PostgreSQL [43]. Данный выбор обусловлен наличием реляционных отношений в описанной системе, а также количеством полей у каждой сущности меньше 10, таким образом, данная СУБД может удовлетворить все потребности при реализации.

В качестве средства разработки метода распознавания суицидальных паттернов поведения человека по текстовым сообщениям использовался ЯП Python. Данный выбор обусловлен следующими факторами:

- большое количество реализаций средств анализа и предобработки текста;
- широкий выбор библиотек для разработки в области машинного обучения;
- просто синтаксиса языка и высокая скорость разработки.

В качестве среды разработки был задействован Visual Studio Code. Данный выбор обусловлен тем, что это ПО распространяется по свободной лицензии, поставляется для конечного пользователя с открытым исходным ко-

дом, а также имеет большое число расширений, ускоряющих процесс разработки.

Список задействованных в разработанном методе библиотек:

- pandas [44] – библиотека для обработки и анализа данных;
- numpy [45] – библиотека, добавляющая поддержку больших многомерных массивов и матриц, вместе с большой библиотекой высокоуровневых математических функций для операций с этими массивами;
- matplotlib [46] – библиотека для визуализации данных;
- scikit-learn [47] – библиотека множества операций и алгоритмов, используемых в сфере науки о данных и машинном обучении;
- nltk [48] – библиотека, предоставляющая обширный набор инструментов для работы с естественными языками;
- pymorphy2 [49] – библиотека, предоставляющая морфологический анализатор, а также утилиты для взаимодействия с ним.

### 3.2 Интерфейс средства сбора данных

На рисунках 3.1-3.4 представлен интерфейс реализованного Telegram-бота.

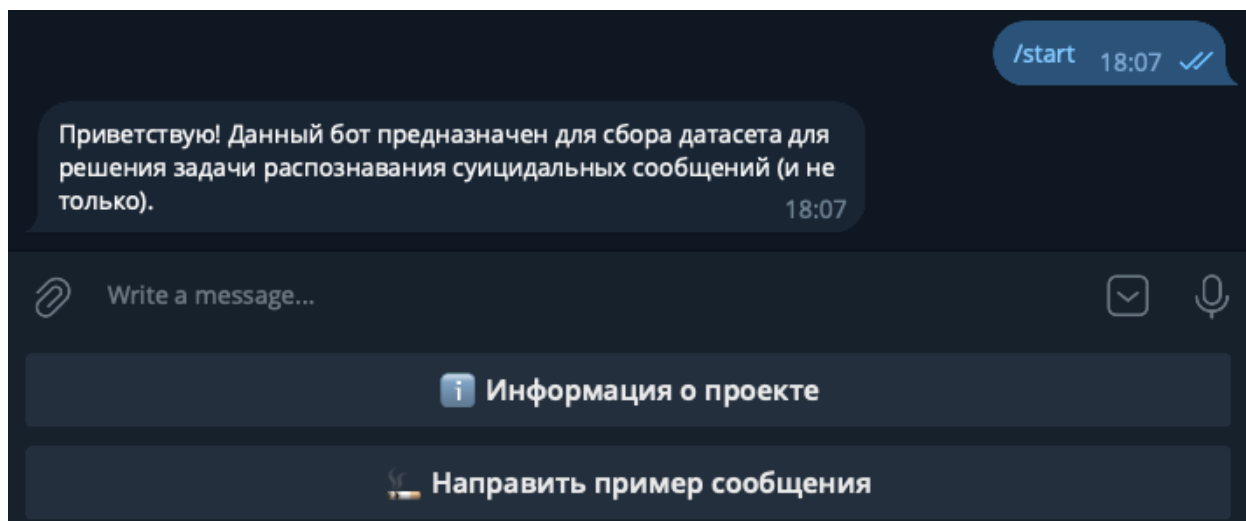


Рисунок 3.1 – Приветственное сообщение новому пользователю.

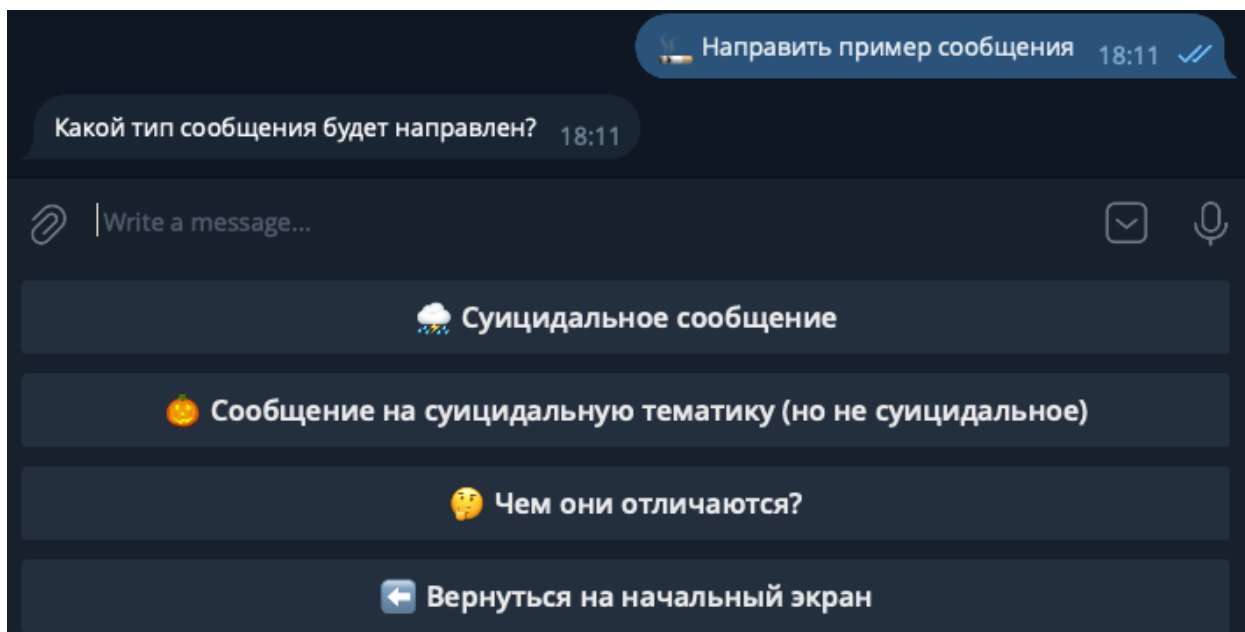


Рисунок 3.2 – Функционал направления в систему сообщения пользователем.

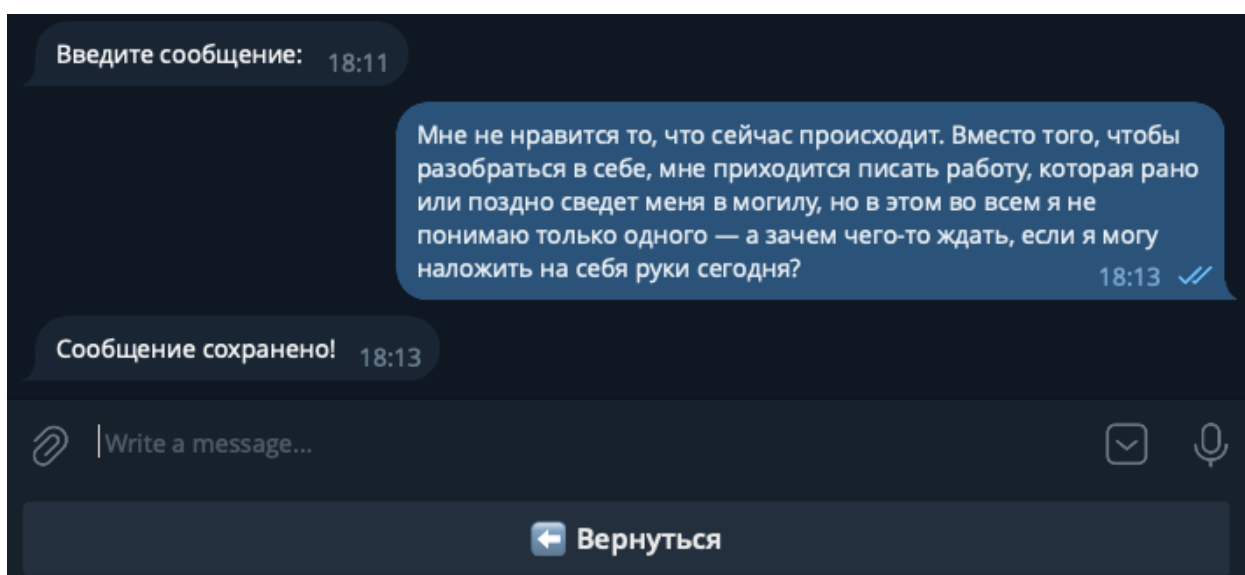


Рисунок 3.3 – Пример результата направленного в систему суицидального сообщения.

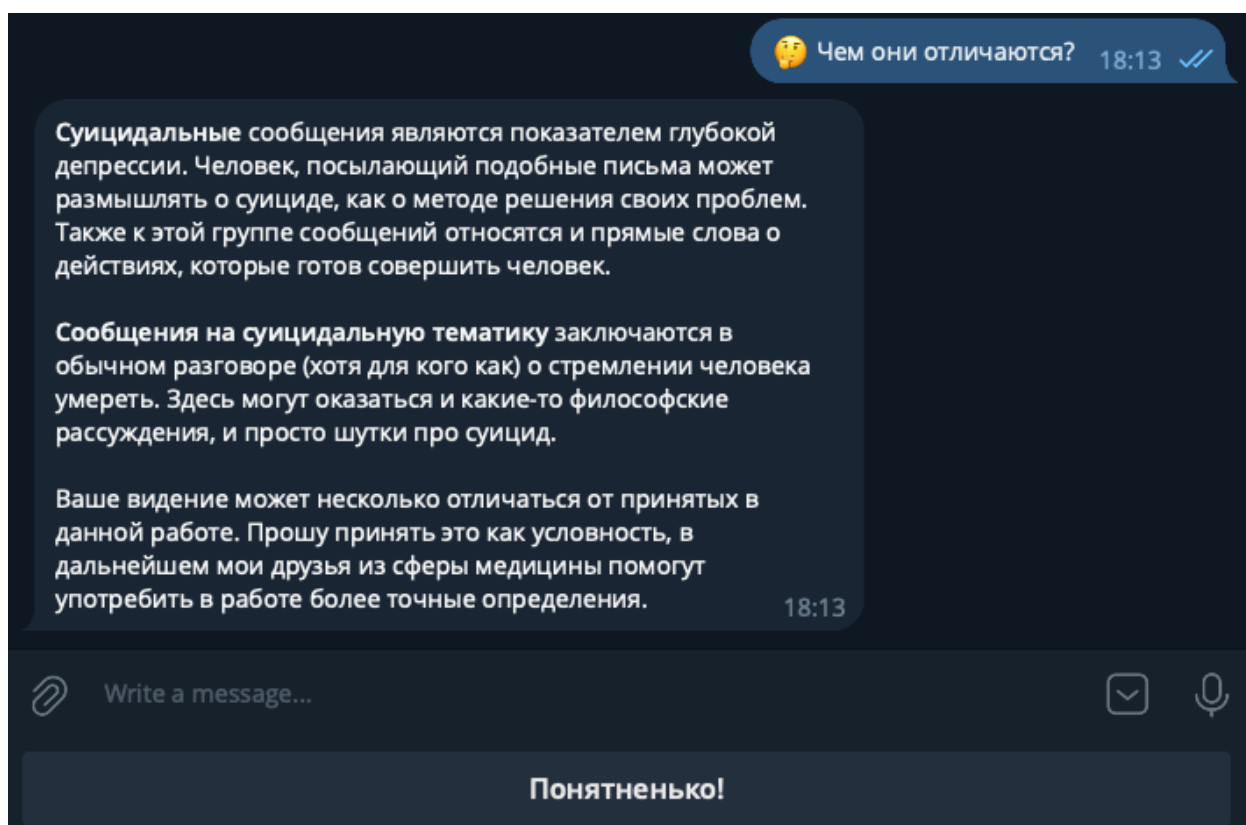


Рисунок 3.4 – Описание отличий суицидального сообщения и сообщения на суицидальную тематику.

### 3.3 Интерфейс средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям

На рисунках 3.5 и 3.6 представлен интерфейс реализованного средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям. Средство позволяет выбрать пользователю как модель, так и метод векторизации сообщения, поступающего в систему.

```
) sh run.sh
Запуск средства распознавания суицидальных паттернов...

Выберите модель, которой хотите воспользоваться.
  1. Градиентный бустинг
  2. Случайный лес
  3. Метод опорных векторов
  4. К-ближайших соседей
  5. Логистическая регрессия
  6. Перцептрон

  7. Выход

Ваш выбор: 2

Какую векторизацию использовать?
  1. Bag of words (Default)
  2. BERT

Ваш выбор: 2

Введите сообщение, которое требуется проанализировать: Увидеть Париж и умереть
Анализ выполняется...
Сообщение не относится к суицидальным

Выберите модель, которой хотите воспользоваться.
  1. Градиентный бустинг
  2. Случайный лес
  3. Метод опорных векторов
  4. К-ближайших соседей
  5. Логистическая регрессия
  6. Перцептрон

  7. Выход

Ваш выбор: 7

Всего доброго!
```

Рисунок 3.5 – Полная пользовательская история использования средства.

```
Выберите модель, которой хотите воспользоваться.
1. Градиентный бустинг
2. Случайный лес
3. Метод опорных векторов
4. К-ближайших соседей
5. Логистическая регрессия
6. Перцептрон

7. Выход

Ваш выбор: 5

Какую векторизацию использовать?
1. Bag of words (Default)
2. BERT

Ваш выбор: 1

Введите сообщение, которое требуется проанализировать: Я хочу умереть, мне все жутко надоело
Анализ выполняется...
Сообщение суицидальное
```

Рисунок 3.6 – Результат анализа суицидального сообщения.

### 3.4 Описание обрабатываемых данных

В результате работы средства сбора данных было размечено 1000 суицидальных сообщений. К собранным сообщениям было добавлено еще 1000 несуйцидальных сообщений из датасета обнаружения пресуйцидальных сигналов [26].

На рисунках 3.7 и 3.8 представлены круговые диаграммы тональности сообщений, полученные с использованием библиотеки Dostoevsky [50].

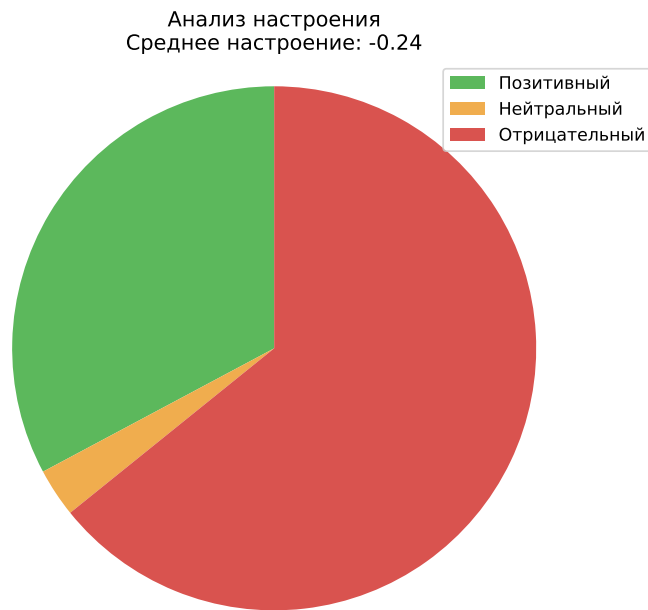


Рисунок 3.7 – Круговая диаграмма тональности суицидальных сообщений.

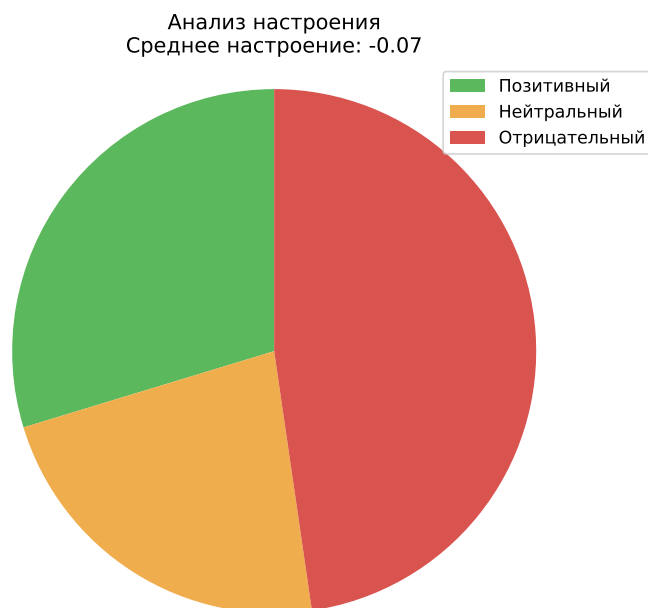


Рисунок 3.8 – Круговая диаграмма тональности несуйцидальных сообщений.

Представленные диаграммы показывают, что практически треть суицидальных сообщений автоматизированное средство оценки тональности распознает как сообщения с отрицательной окраской. Однако наличие среди

них позитивно настроенных сообщений – ошибка распознавания модели. Среди несуицидальных сообщений преобладают тексты с отрицательной окраской, однако тут их уже меньше половины, а нейтральных сообщений почти что четверть из всех представленных.

На рисунке 3.9 представлена визуализация собранных данных класса суицидальных сообщений. Чаще всего в суицидальных сообщениях фигурируют слова “жизнь” (585 раз), “хотеть” (556 раз), “человек” (491 раз) и “мочь” (452 раза). Также стоит обратить внимание на присутствие слов “суицид”, “страдать”, “депрессия”, “смерть”, “умирать” и “ад”.



Рисунок 3.9 – Облако слов класса суицидальных сообщений.

На рисунке 3.10 представлена визуализация данных класса несуицидальных сообщений. Чаще всего в несуицидальных сообщениях встречаются слова “хотеть” (159 раз) и “человек” (67 раз). Кроме того сообщения данной тематики чаще включают в себя различные вариации нецензурной брани.



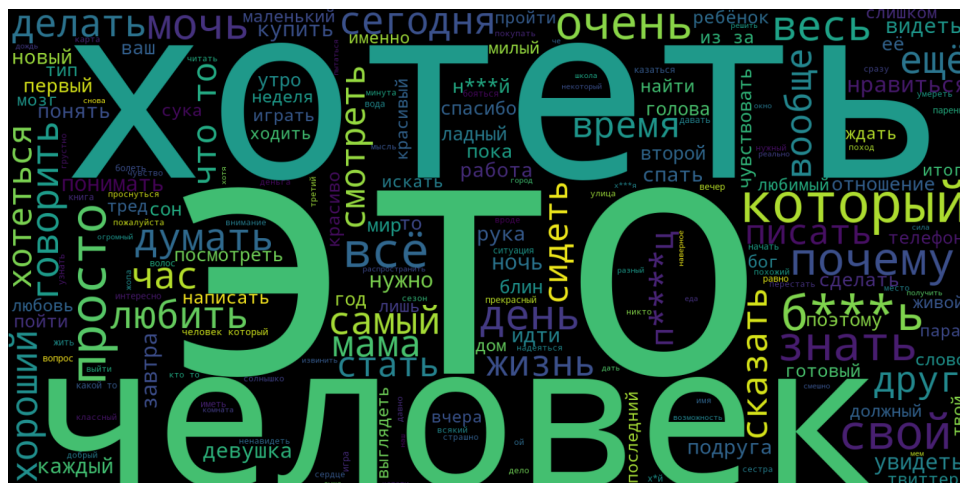


Рисунок 3.10 – Облако слов класса несуицидальных сообщений.

Представленная информация подтверждает факт того, что выбранные классы разделимы и отличны частотой употребления как слов, так и тематик. Кроме того, стоит отметить, что слово “хотеть” встречается в суицидальных сообщениях в  $\approx 7.83$  раза чаще, чем в несуйцидальных, а слово “человек” – в  $\approx 7.33$  раза чаще. Таким образом, суицидальные сообщения являются менее “разнообразными” и фиксирующимися на определенном словарном множестве.

## Вывод

В качестве инструментов разработки средства сбора данных будут задействованы ЯП Python и библиотека Telebot. В качестве языка разработки средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям также будет использоваться ЯП Python. Задействованные библиотеки: pandas, numpy, matplotlib, scikit-learn, nltk, pymorphy2.

Были представлены интерфейсы средства сбора данных и средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям.

Представленные диаграммы тональности сообщений показали, что практически треть суицидальных сообщений автоматизированное средство оценки тональности распознает как сообщения с отрицательной окраской. Среди несуйцидальных сообщений преобладают тексты с отрицательной окраской, при этом нейтральных сообщений – четверть из всех.

Визуализированные облака слов подтвердили гипотезу, что выбранные классы суицидальных и несуйцидальных сообщений разделимы и отличны частотой некоторых слов. Отмечено, что слово “хотеть” встречается в суицидальных сообщениях в  $\approx 7.83$  раза чаще, чем в несуйцидальных, а слово “человек” – в  $\approx 7.33$  раза чаще. Таким образом, суицидальные сообщения являются менее “разнообразными” и фиксирующимися на определенном словарном множестве.

## **4. Исследовательский раздел**

### **4.1 Условия исследований**

Исследование проводилось на персональном компьютере со следующими характеристиками:

- процессор Apple M1 Pro,
- операционная система macOS Ventura 13.5.2 (22G91),
- 32 Гб оперативной памяти.

Для определения гиперпараметров каждой модели применялся метод поиска по сетке с опорой на значение F1-меры. Разбиение данных на выборки производилось на 4 части, 1 из которых используется в качестве тестовой, результаты приводятся для каждого разбиения.

### **4.2 Исследование применимости моделей машинного обучения в методе распознавания суицидальных паттернов поведения человека по текстовым сообщениям**

#### **4.2.1 Градиентный бустинг**

Параметры модели при применении метода векторизации “мешок слов”:

- коэффициент обучения – влияет на скорость обучения – 0.5;
- количество классификаторов в ансамбле – определяет количество деревьев, используемых в алгоритме – 215;
- минимальное количество образцов для разбиения узла дерева решений – 5.

На рисунке 4.1 представлены матрицы ошибок, полученные с использованием градиентного бустинга, метод векторизации – “мешок слов”.

На рисунке 4.2 представлены оценки классификатора, полученные с использованием градиентного бустинга, метод векторизации – “мешок слов”.

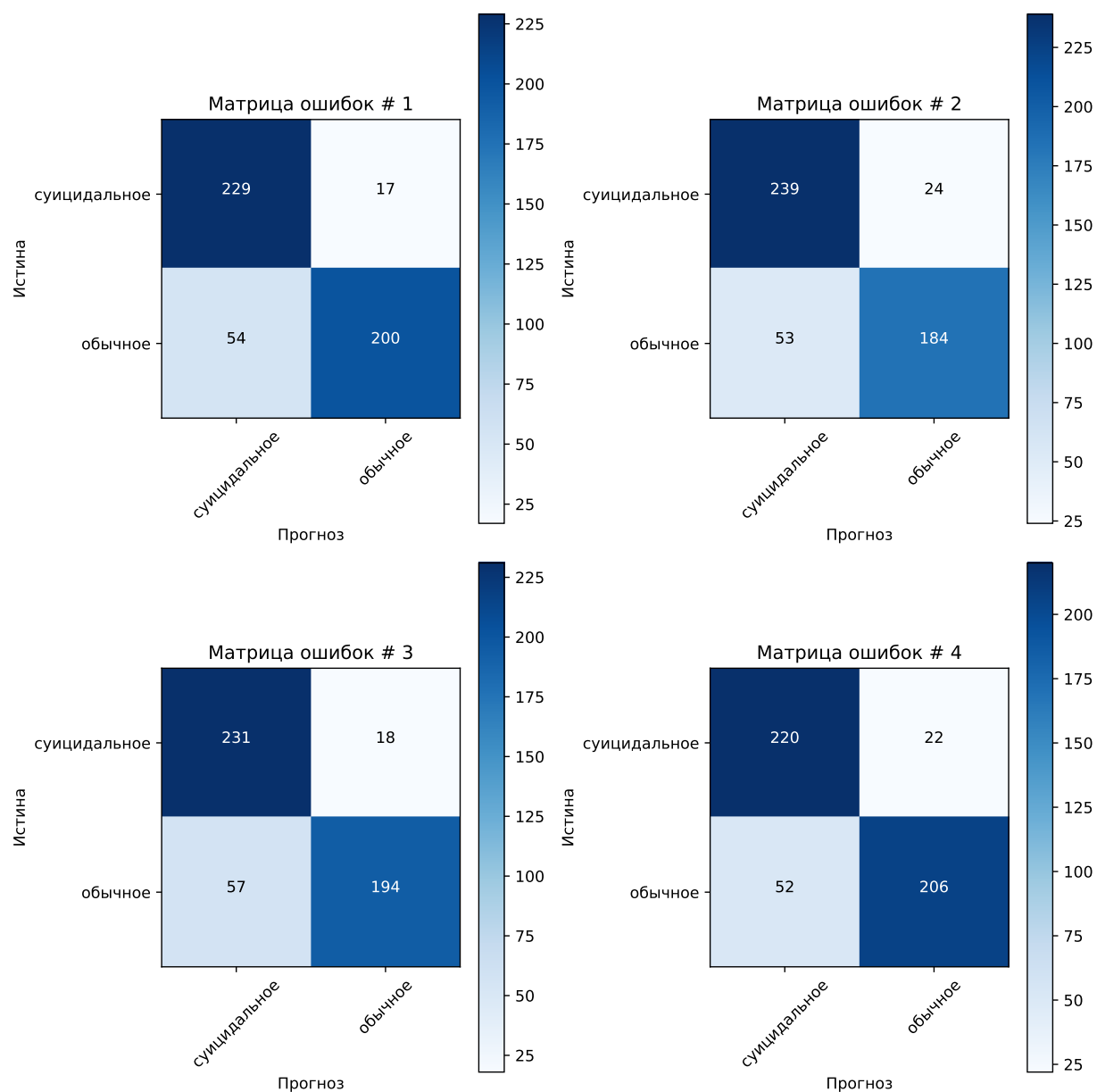


Рисунок 4.1 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием градиентного бустинга (метод векторизации – “мешок слов”).

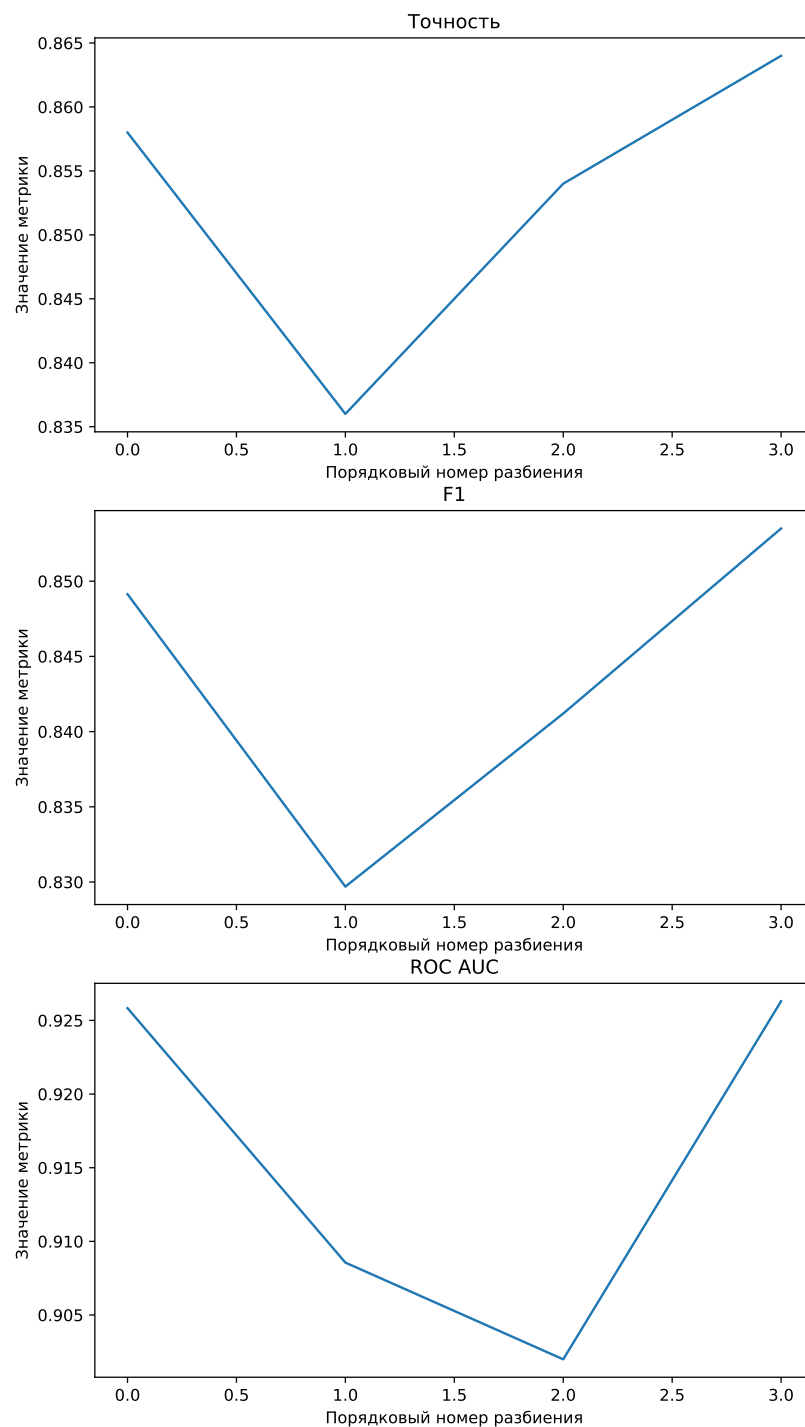


Рисунок 4.2 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием градиентного бустинга (метод векторизации – “мешок слов”).

Параметры модели при применении векторизации BERT:

- коэффициент обучения – 0.3;
- количество классификаторов в ансамбле – 245;
- минимальное количество образцов для разбиения узла дерева решений – 16.

На рисунке 4.3 представлены матрицы ошибок, полученные с использованием градиентного бустинга, метод векторизации – BERT.

На рисунке 4.4 представлены оценки классификатора, полученные с использованием градиентного бустинга, метод векторизации – BERT.

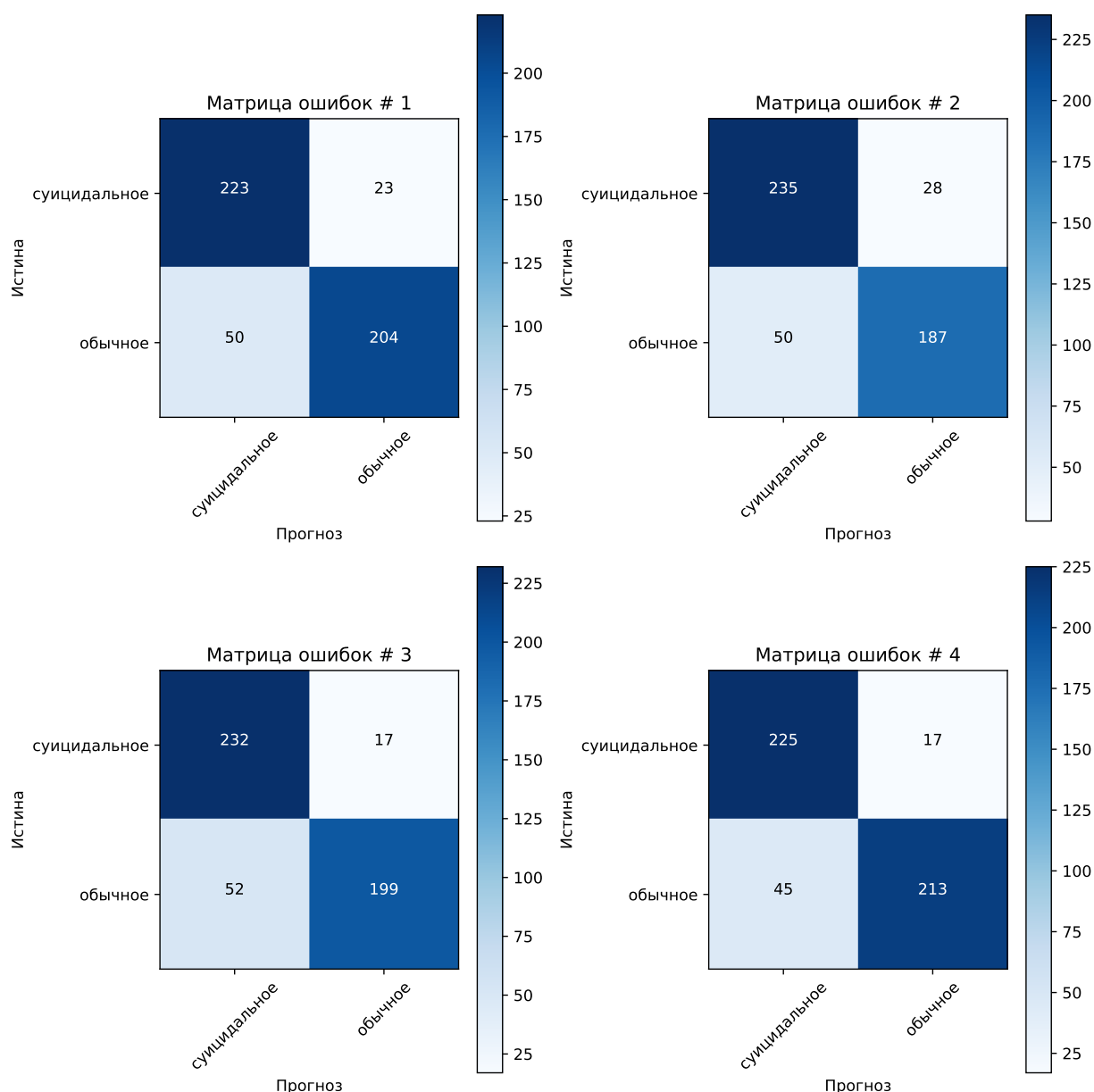


Рисунок 4.3 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием градиентного бустинга (метод векторизации – BERT).

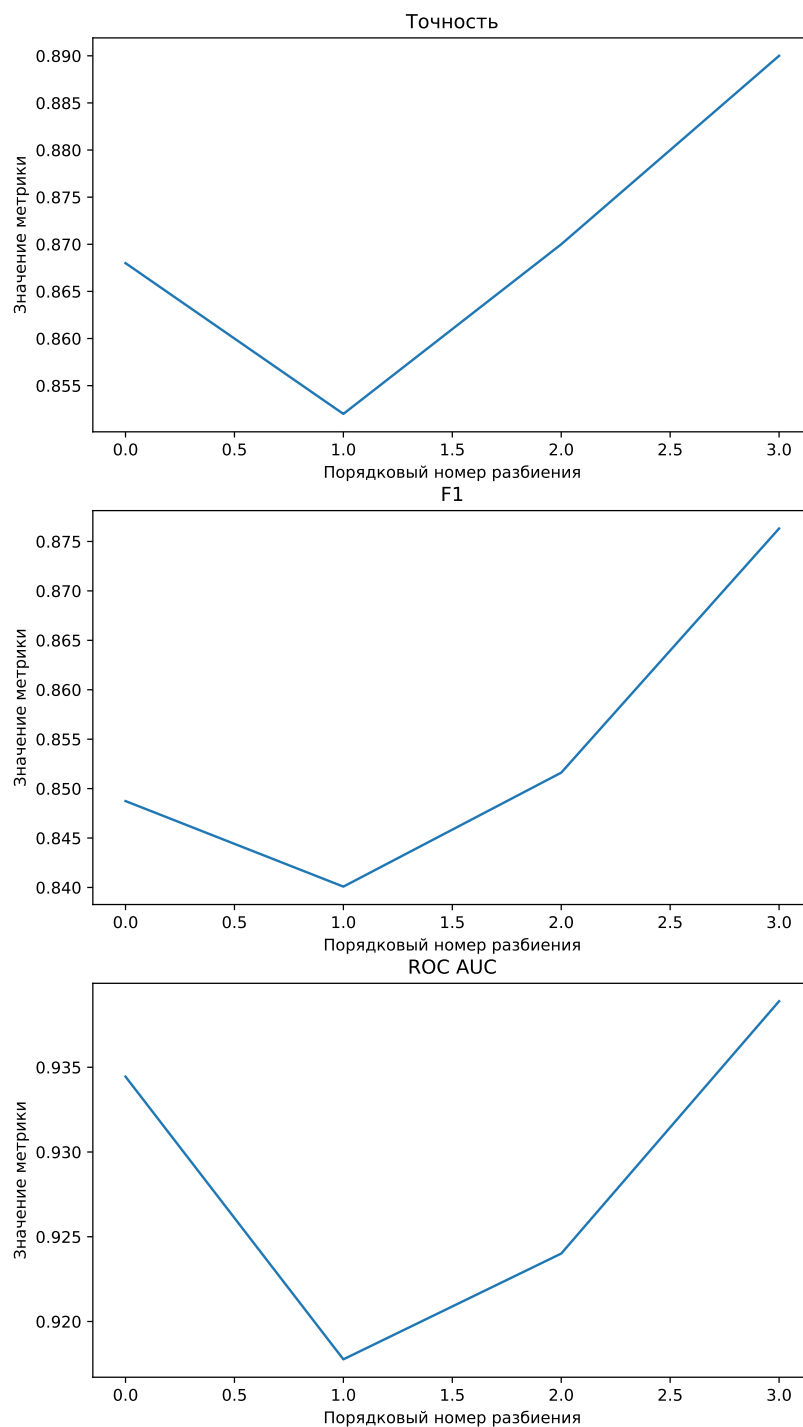


Рисунок 4.4 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием градиентного бустинга (метод векторизации – BERT).

### 4.2.2 Метод случайного леса

Параметры модели при применении метода векторизации “мешок слов”:

- максимальная глубина – определяет максимальную глубину дерева решений – 300;
- максимальное количество признаков – определяет максимальное количество признаков, используемых для разбиения узла дерева решений, может принимать значение функции –  $\log_2$ ;
- количество классификаторов в ансамбле – определяет количество деревьев, используемых в алгоритме, – 850.

На рисунке 4.5 представлены матрицы ошибок, полученные с использованием метода случайного леса, метод векторизации – “мешок слов”.

На рисунке 4.6 представлены оценки классификатора, полученные с использованием метода случайного леса, метод векторизации – “мешок слов”.



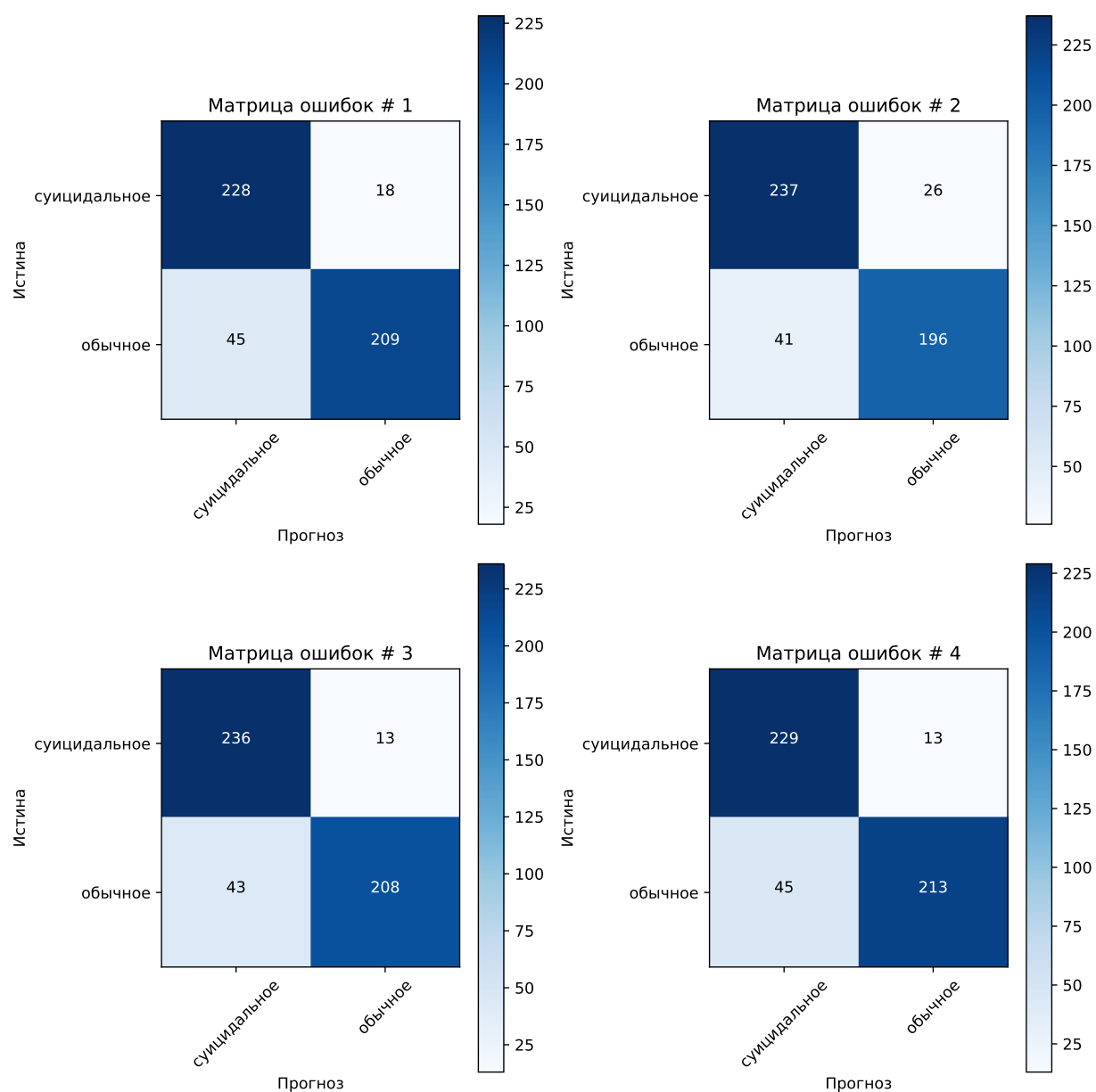


Рисунок 4.5 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием метода случайного леса (метод векторизации – “мешок слов”).

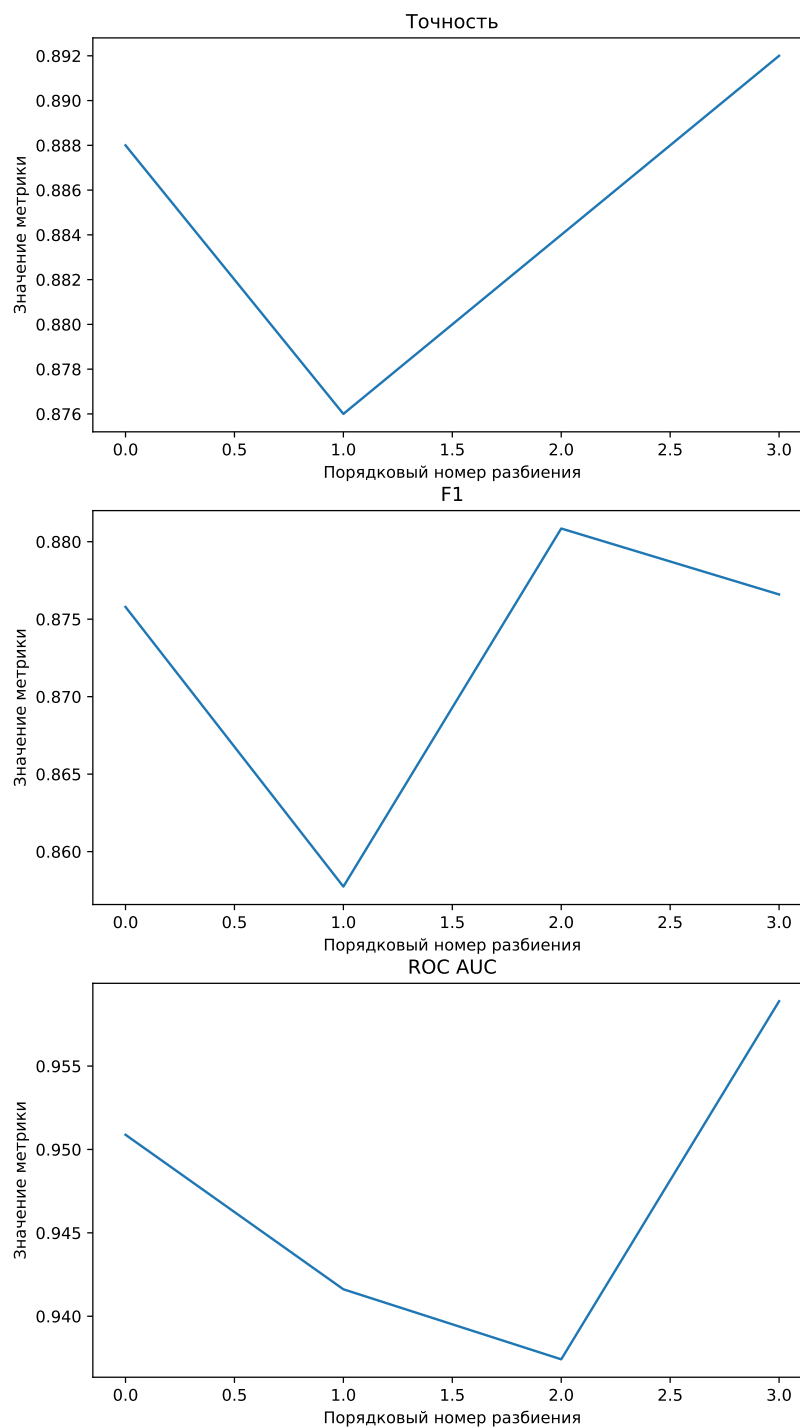


Рисунок 4.6 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием метода случайного леса (метод векторизации – “мешок слов”).

Параметры модели при применении векторизации BERT:

- максимальная глубина – отсутствует;
- максимальное количество признаков –  $\log_2$ ;
- количество классификаторов в ансамбле – 2600.

На рисунке 4.7 представлены матрицы ошибок, полученные с использованием метода случайного леса, метод векторизации – BERT.

На рисунке 4.8 представлены оценки классификатора, полученные с использованием метода случайного леса, метод векторизации – BERT.

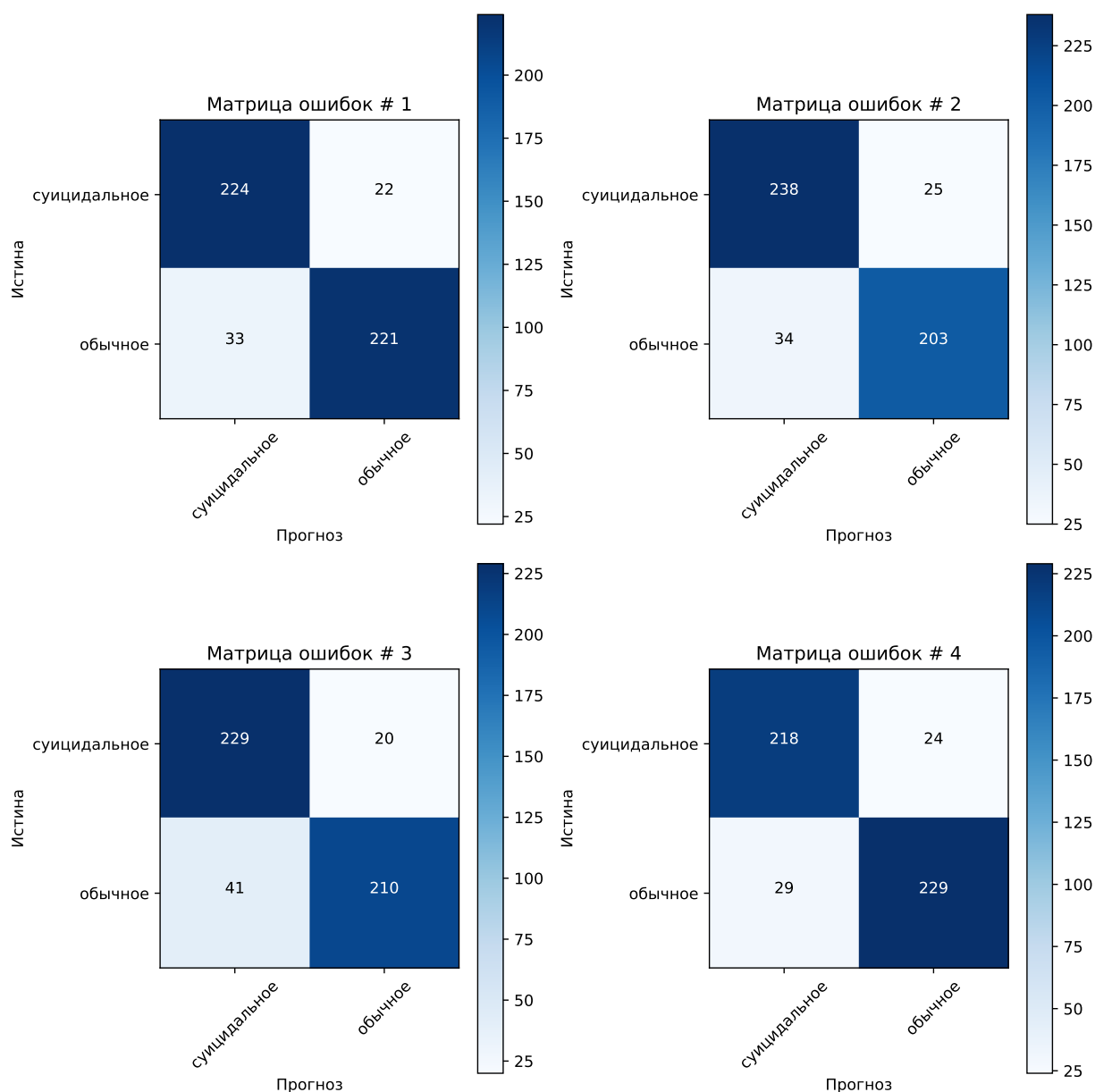


Рисунок 4.7 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием метода случайного леса (метод векторизации – BERT).

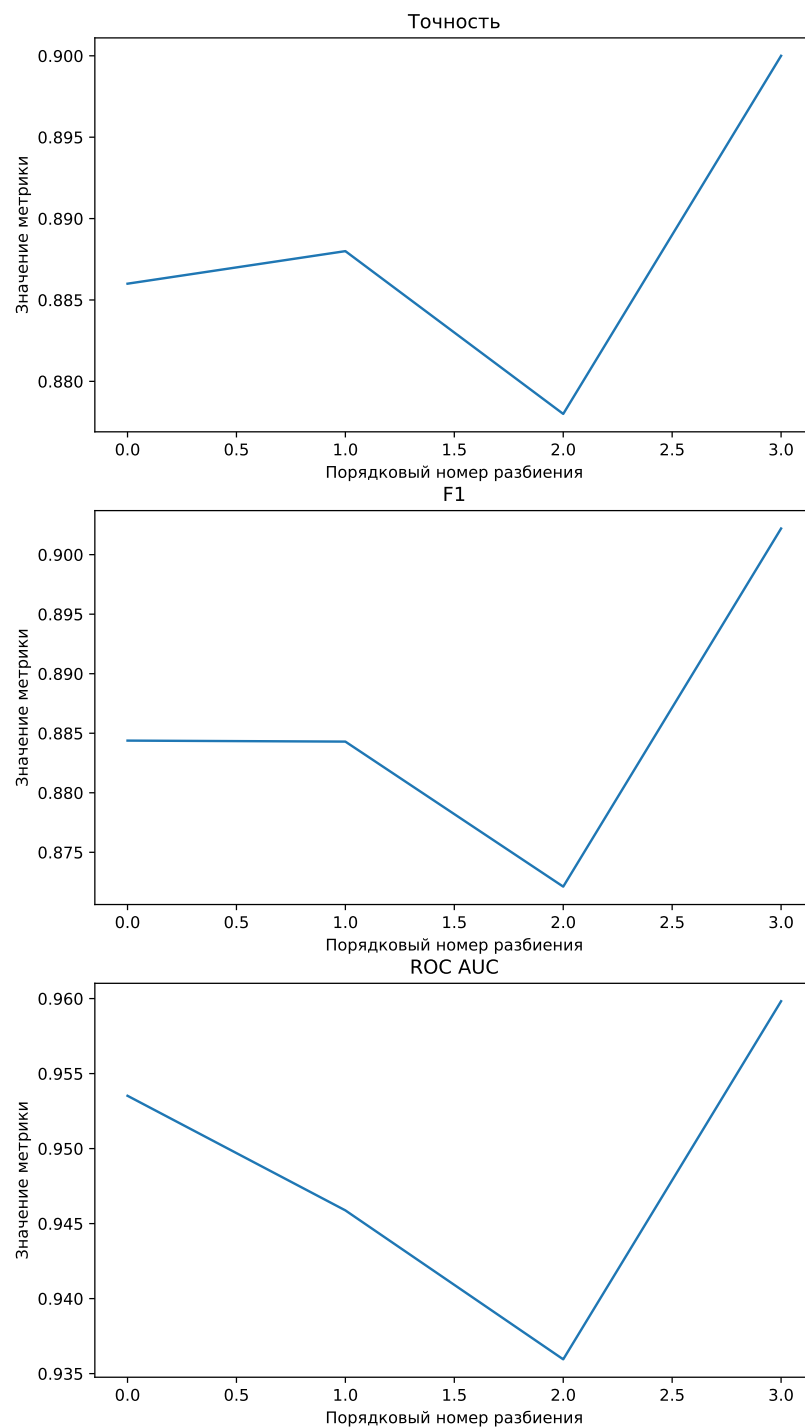


Рисунок 4.8 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием метода случайного леса (метод векторизации – BERT).

### 4.2.3 Метод опорных векторов

Параметры модели при применении метода векторизации “мешок слов”:

- регуляризационный параметр – определяет степень регуляризации модели – 7.0;
- ядро – определяет тип ядра, используемого в модели – RBF (радиальное базисной функции);
- коэффициент ядра – определяет “ширину” ядра – значение вычисляется как обратная величина произведения количества признаков на дисперсию признаков.

На рисунке 4.9 представлены матрицы ошибок, полученные с использованием метода опорных векторов, метод векторизации – “мешок слов”.

На рисунке 4.10 представлены оценки классификатора, полученные с использованием метода опорных векторов, метод векторизации – “мешок слов”.

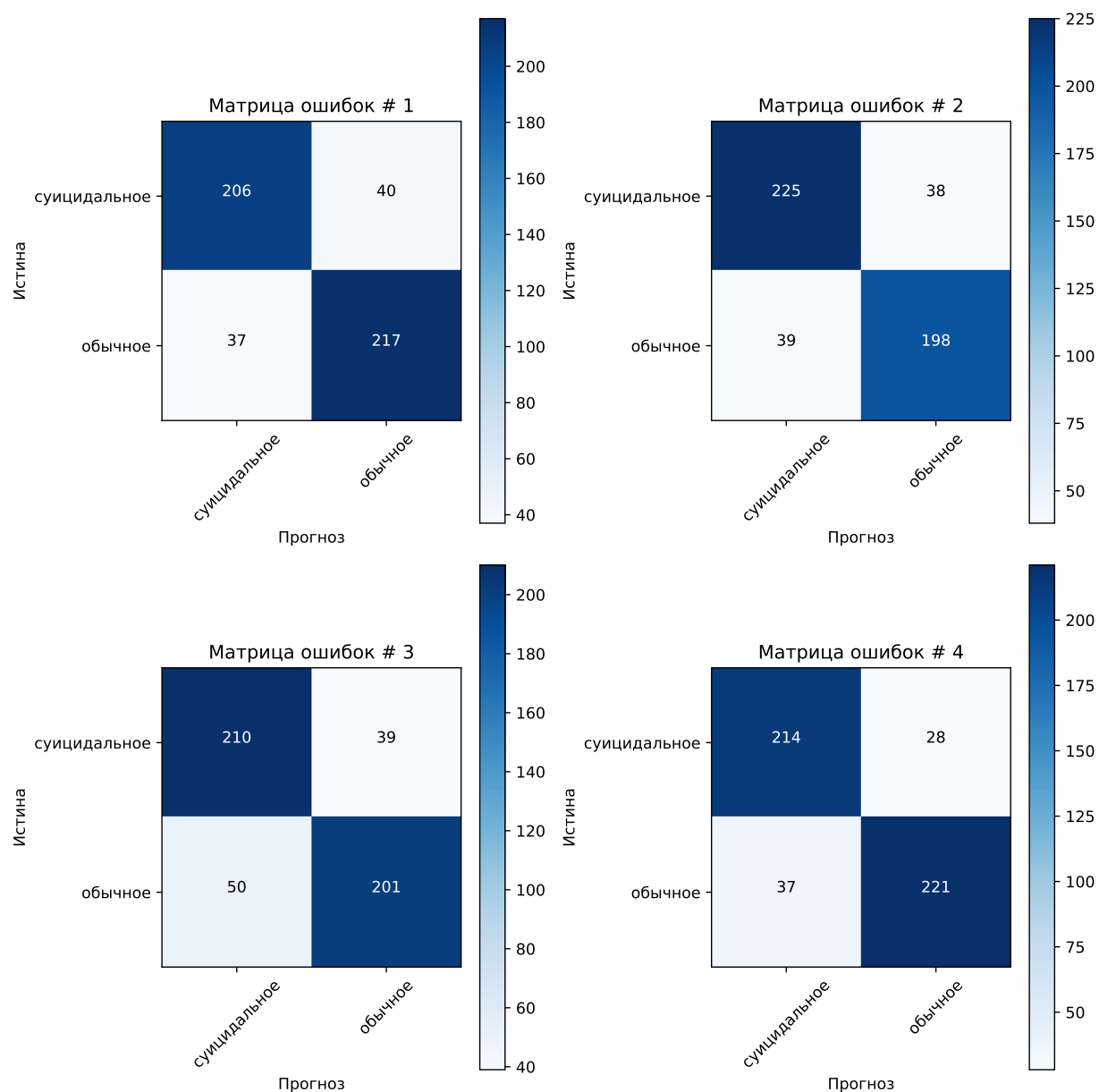


Рисунок 4.9 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием метода опорных векторов (метод векторизации – “мешок слов”).

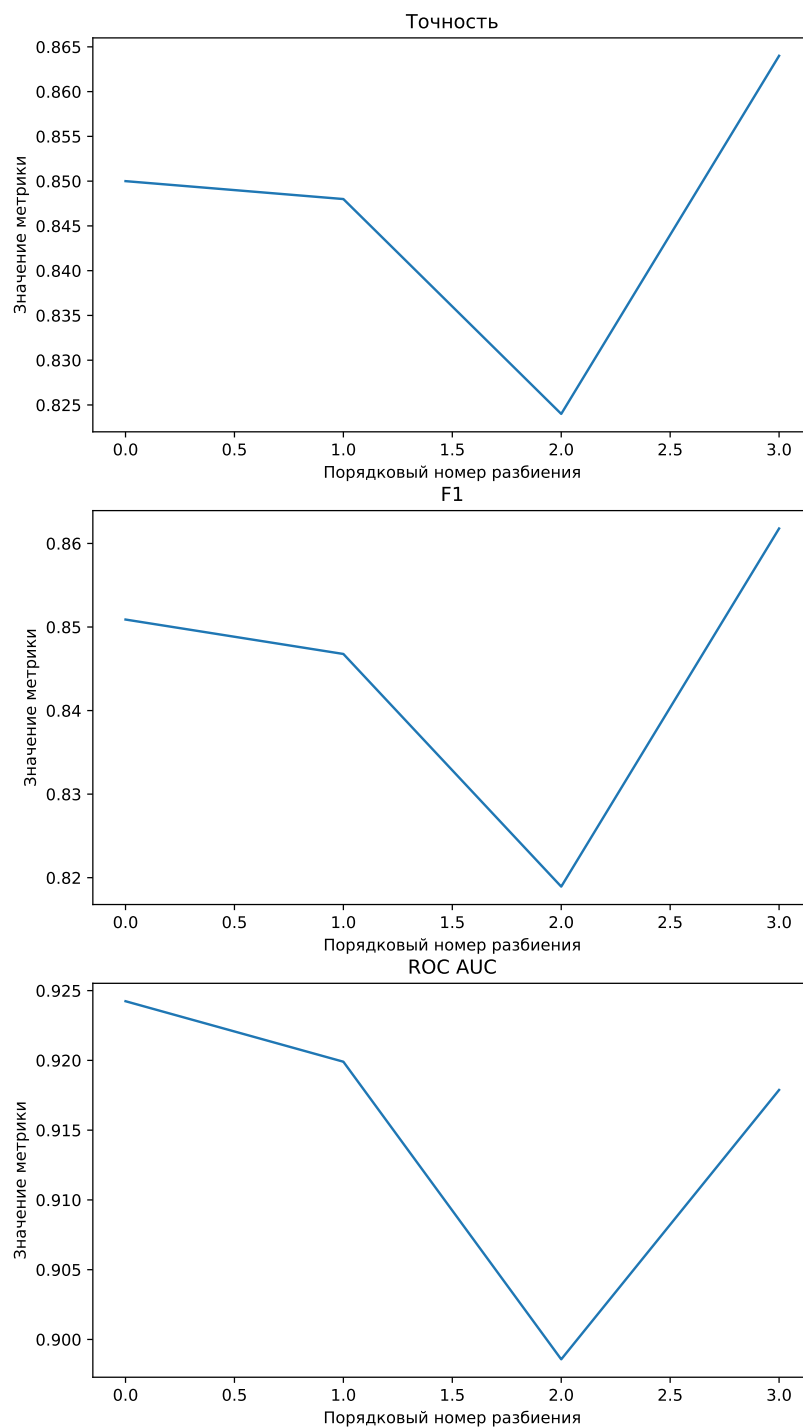


Рисунок 4.10 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием метода опорных векторов (метод векторизации – “мешок слов”).

Параметры модели при применении векторизации BERT:

- регуляризационный параметр – 1.5;
- ядро – линейное;
- коэффициент ядра – значение вычисляется как обратная величина произведения количества признаков на дисперсию признаков.

На рисунке 4.11 представлены матрицы ошибок, полученные с использованием метода опорных векторов, метод векторизации – BERT.

На рисунке 4.12 представлены оценки классификатора, полученные с использованием метода опорных векторов, метод векторизации – BERT.

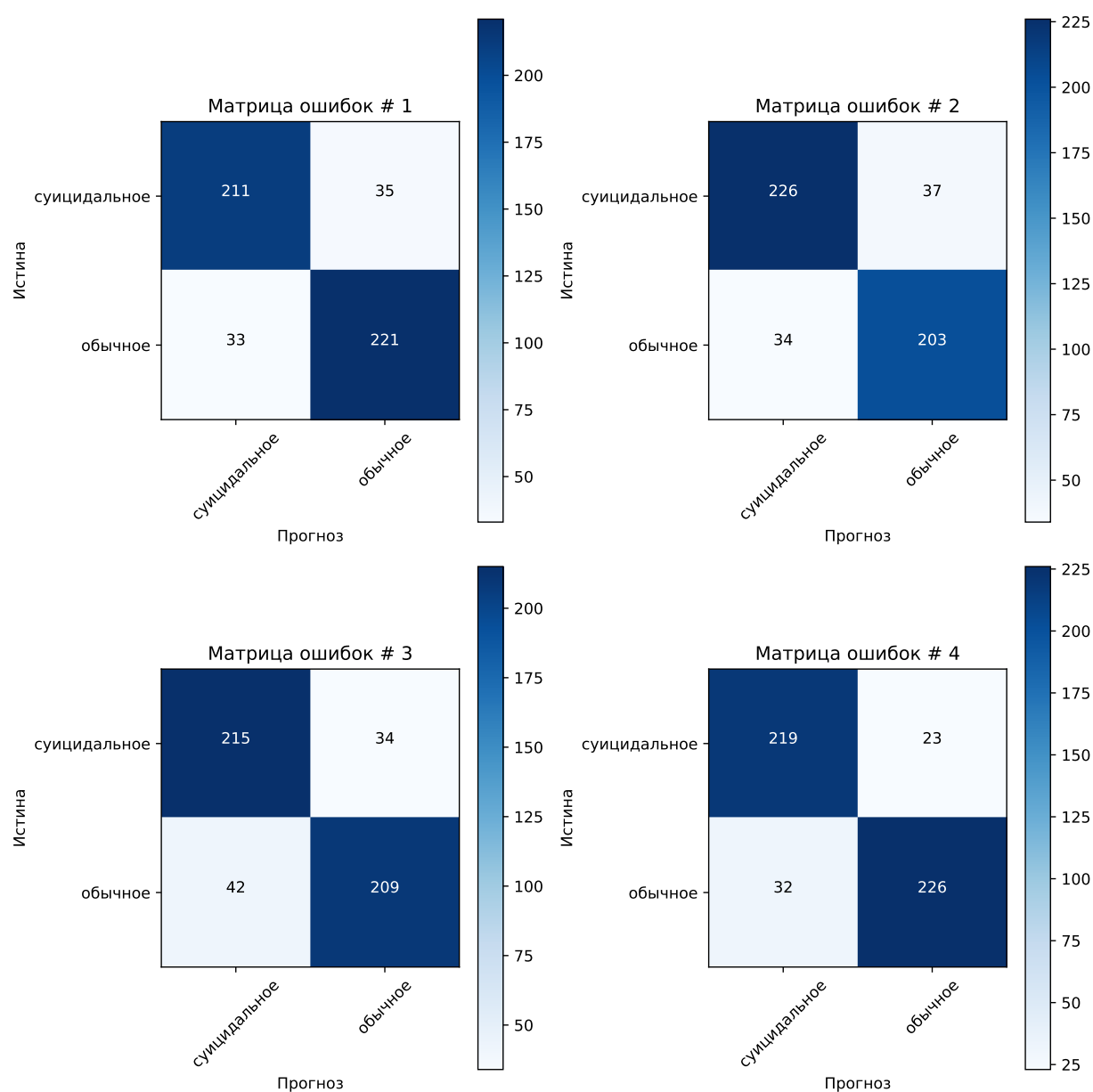


Рисунок 4.11 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием метода опорных векторов (метод векторизации – BERT).



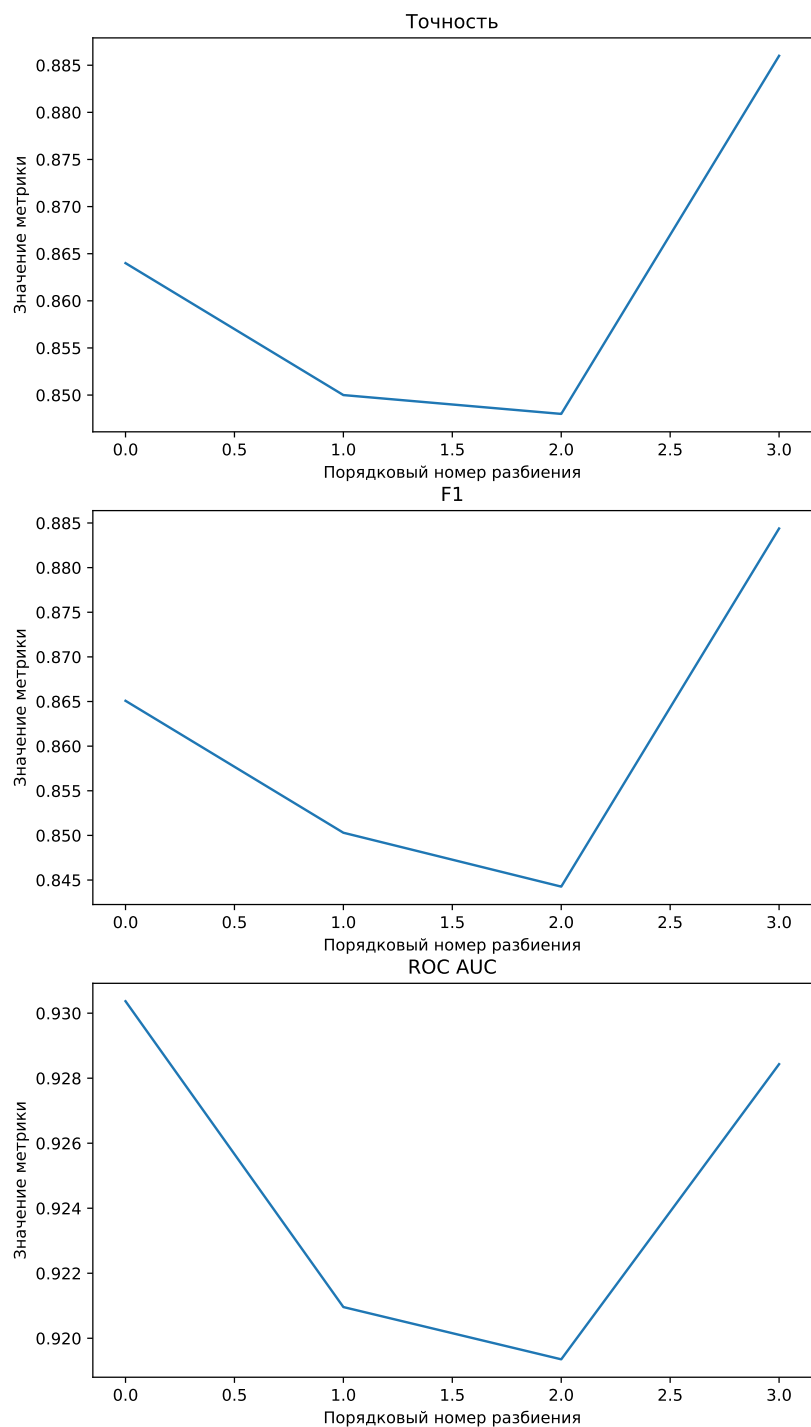


Рисунок 4.12 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием метода опорных векторов (метод векторизации – BERT).

#### 4.2.4 Метод К-ближайших соседей

Параметры модели при применении метода векторизации “мешок слов”:

- количество образцов в листе дерева – 4;
- метрика расстояния – евклидово расстояние;
- количество ближайших соседей – 5;
- метод вычисления весов ближайших соседей – веса, обратно пропорциональные расстоянию до соседа.

На рисунке 4.13 представлены матрицы ошибок, полученные с использованием метода К-ближайших соседей, метод векторизации – “мешок слов”.

На рисунке 4.14 представлены оценки классификатора, полученные с использованием метода К-ближайших соседей, метод векторизации – “мешок слов”.

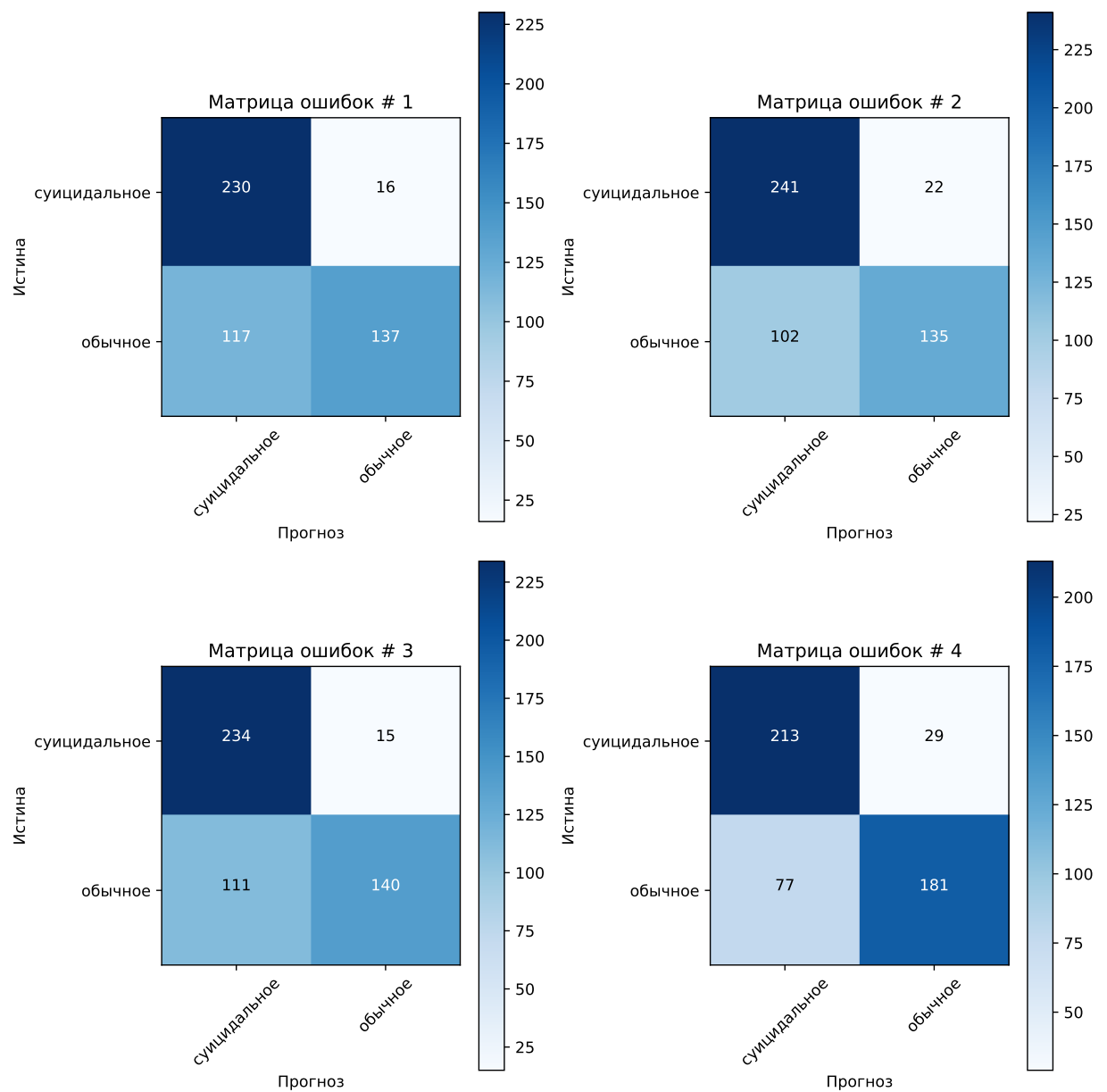


Рисунок 4.13 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием метода К-ближайших соседей (метод векторизации – “мешок слов”).

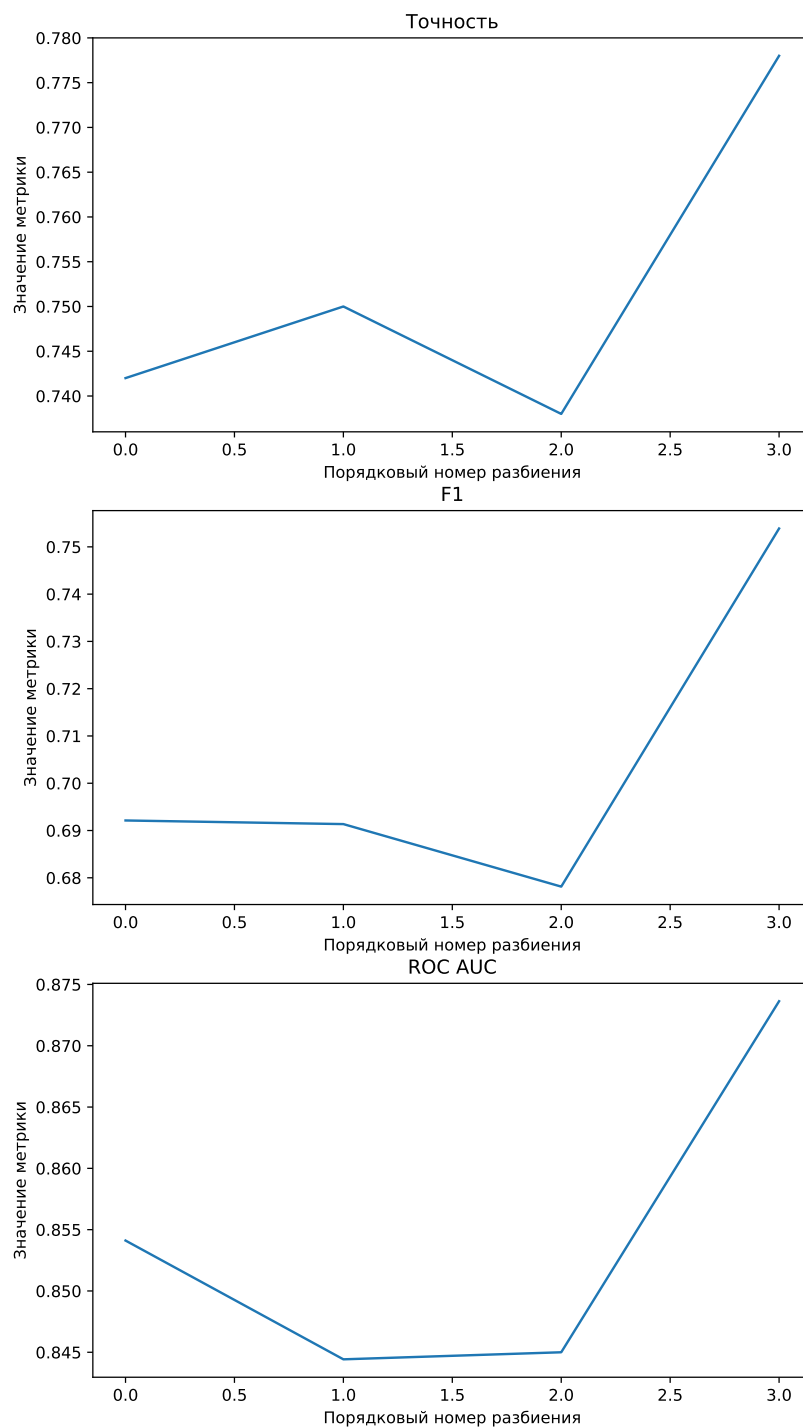


Рисунок 4.14 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием метода К-ближайших соседей (метод векторизации – “мешок слов”).

Параметры модели при применении векторизации BERT:

- количество образцов в листе дерева – 1;
- метрика расстояния – евклидово расстояние;
- количество ближайших соседей – 6;
- метод вычисления весов ближайших соседей – веса, обратно пропорциональные расстоянию до соседа.

На рисунке 4.15 представлены матрицы ошибок, полученные с использованием метода К-ближайших соседей, метод векторизации – BERT.

На рисунке 4.16 представлены оценки классификатора, полученные с использованием метода К-ближайших соседей, метод векторизации – BERT.

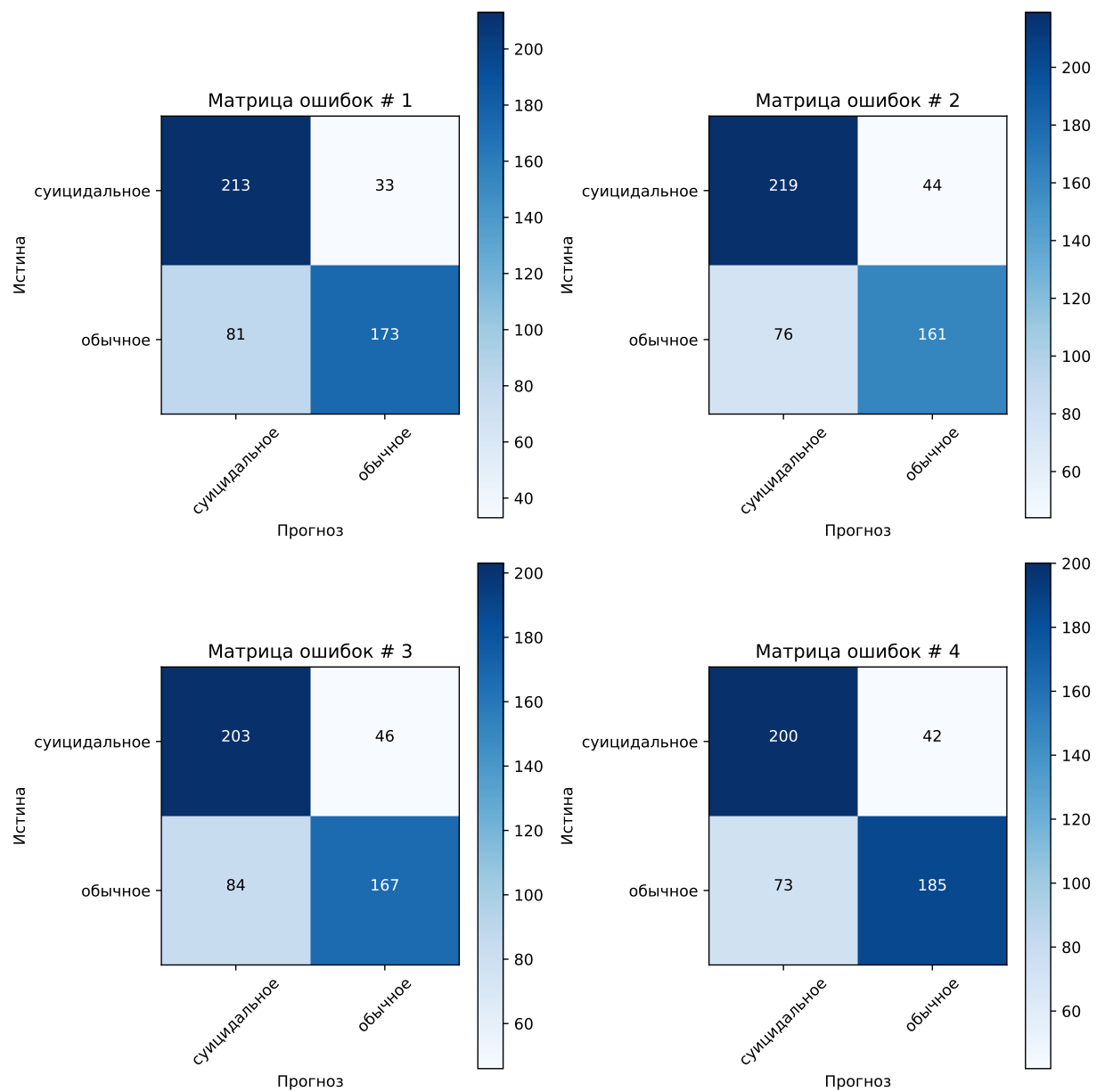


Рисунок 4.15 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием метода К-ближайших соседей (метод векторизации – BERT).

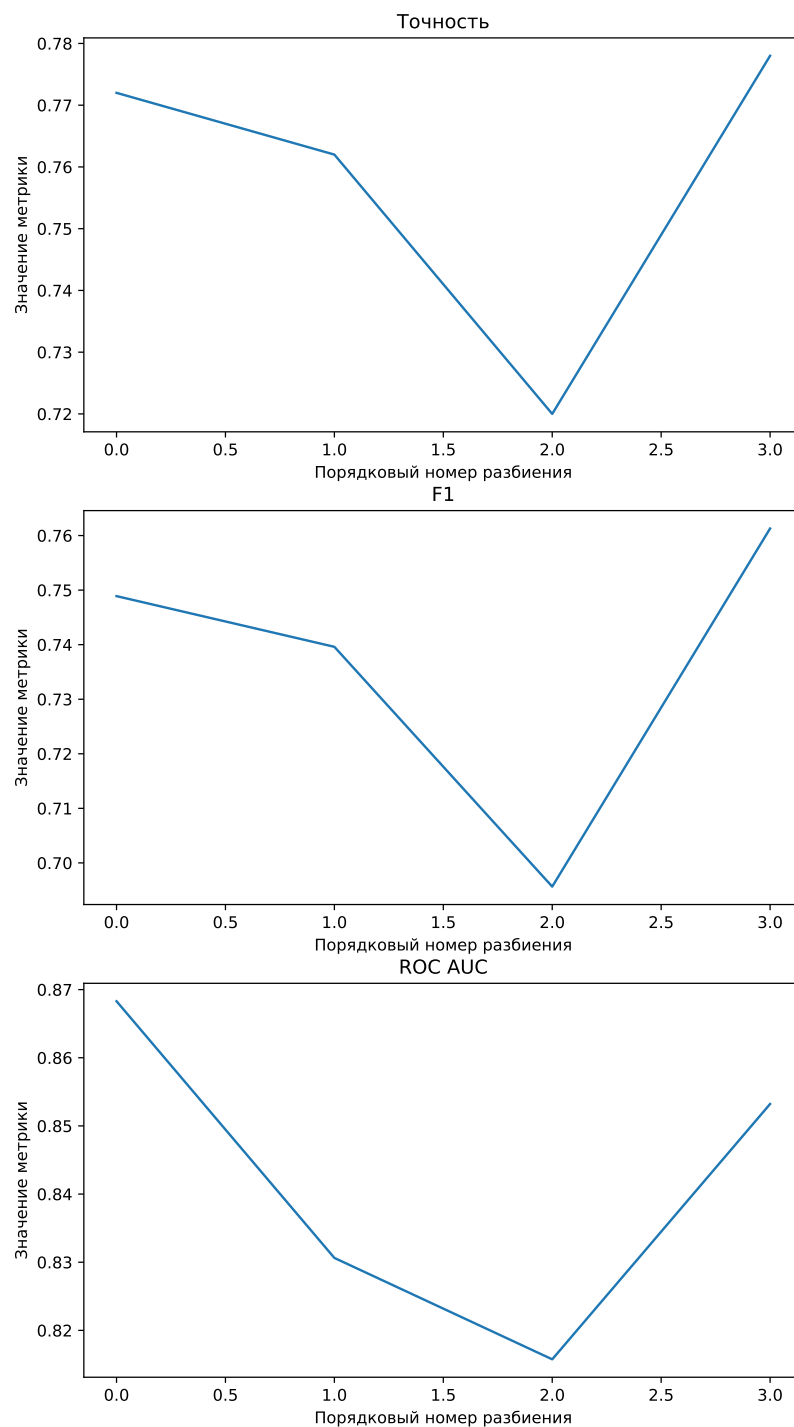


Рисунок 4.16 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием метода К-ближайших соседей (метод векторизации – BERT).

#### 4.2.5 Логистическая регрессия

Параметры модели при применении метода векторизации “мешок слов”:

- регуляризационный параметр – 2.2;
- веса классов – сбалансированно;
- штраф – определяет тип штрафа, используемого для регуляризации – сумма квадратов весов модели, умноженных на гиперпараметр регуляризации;
- алгоритм оптимизации – liblinear.

На рисунке 4.17 представлены матрицы ошибок, полученные с использованием логистической регрессии, метод векторизации – “мешок слов”.

На рисунке 4.18 представлены оценки классификатора, полученные с использованием логистической регрессии, метод векторизации – “мешок слов”.



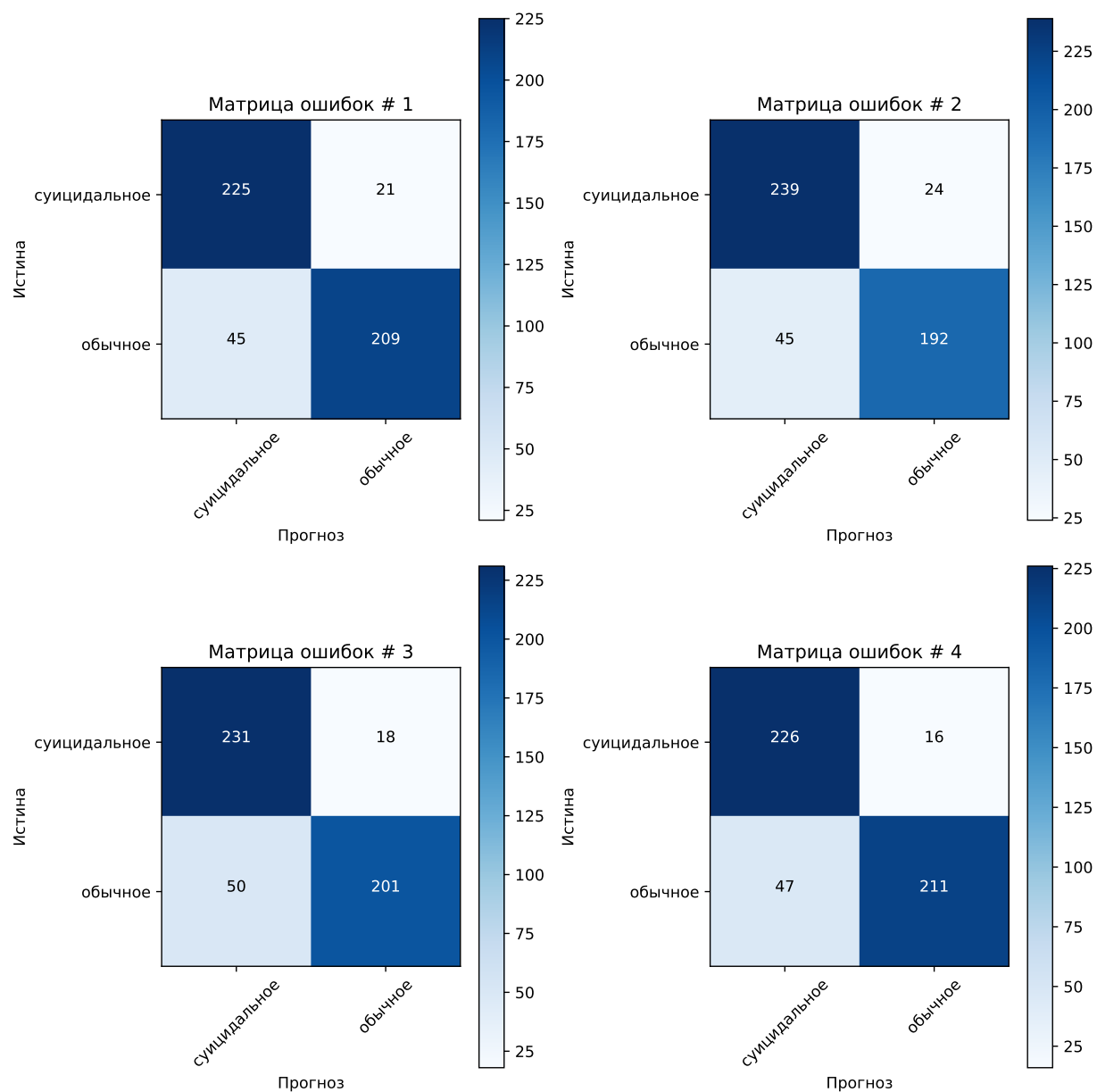


Рисунок 4.17 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием логистической регрессии (метод векторизации – “мешок слов”).

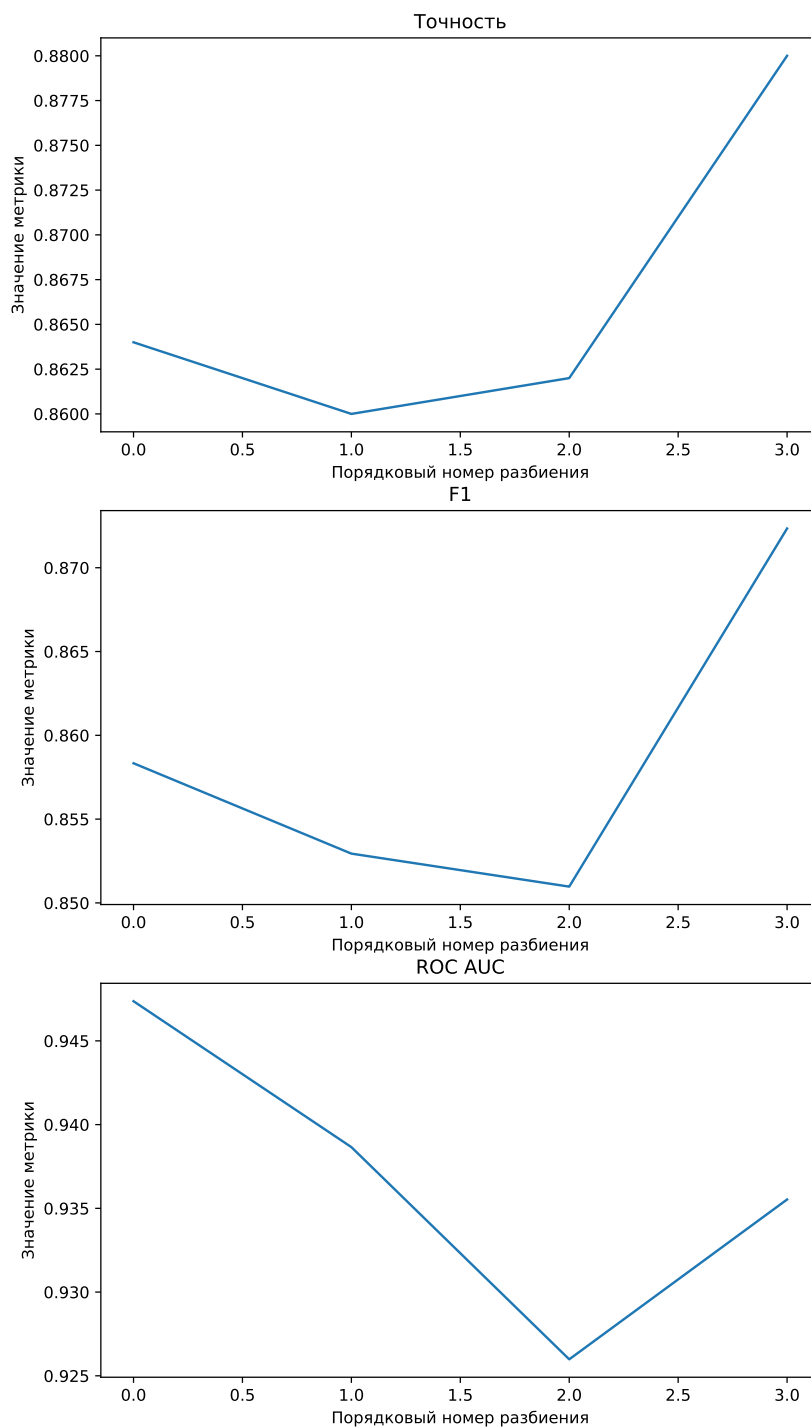


Рисунок 4.18 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием логистической регрессии (метод векторизации – “мешок слов”).

Параметры модели при применении векторизации BERT:

- регуляризационный параметр – 2.1;
- веса классов – сбалансированно;
- штраф – сумма квадратов весов модели, умноженных на гиперпараметр регуляризации;
- алгоритм оптимизации – liblinear.

На рисунке 4.19 представлены матрицы ошибок, полученные с использованием логистической регрессии, метод векторизации – BERT.

На рисунке 4.20 представлены оценки классификатора, полученные с использованием логистической регрессии, метод векторизации – BERT.

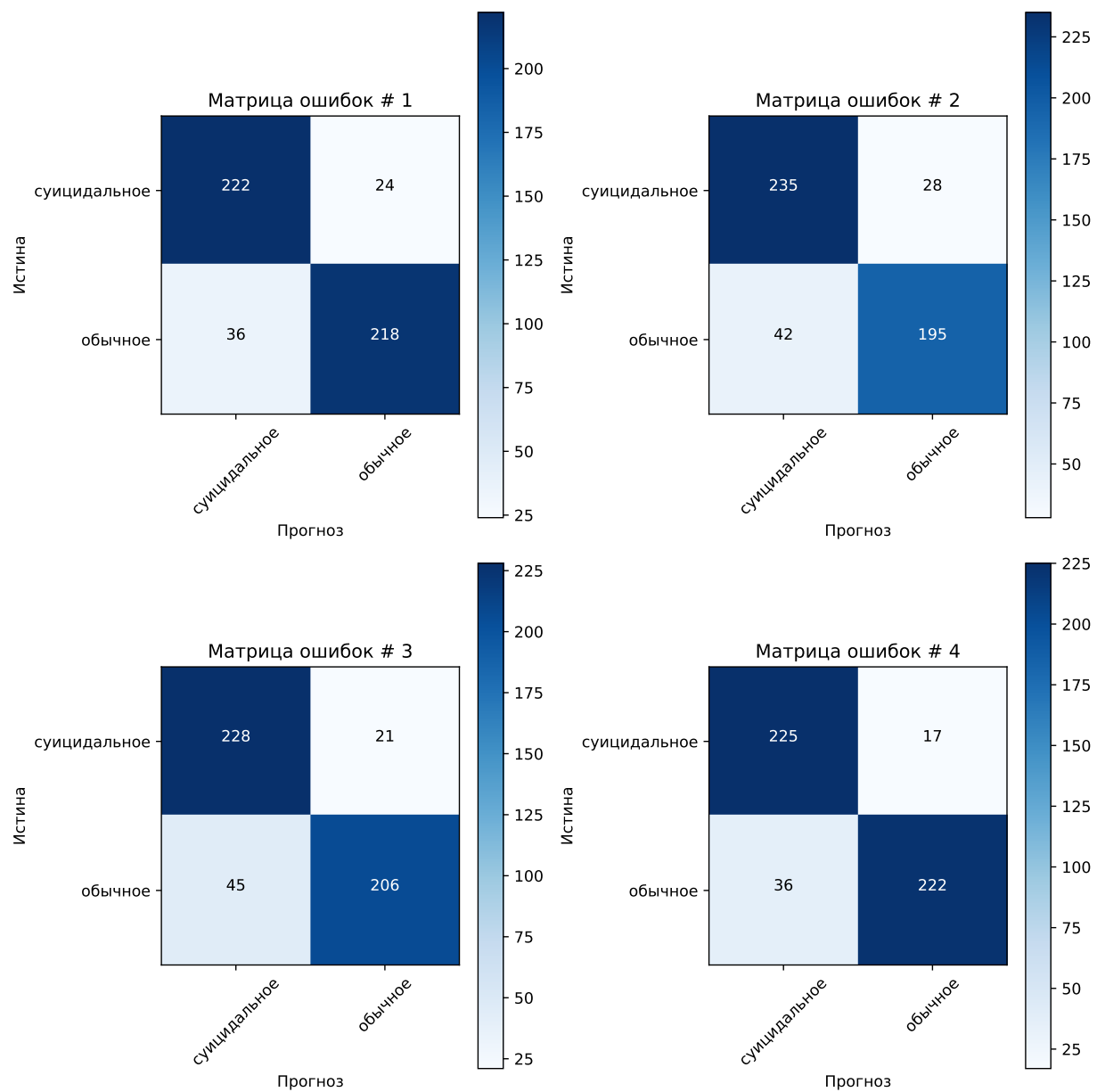


Рисунок 4.19 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием логистической регрессии (метод векторизации – BERT).

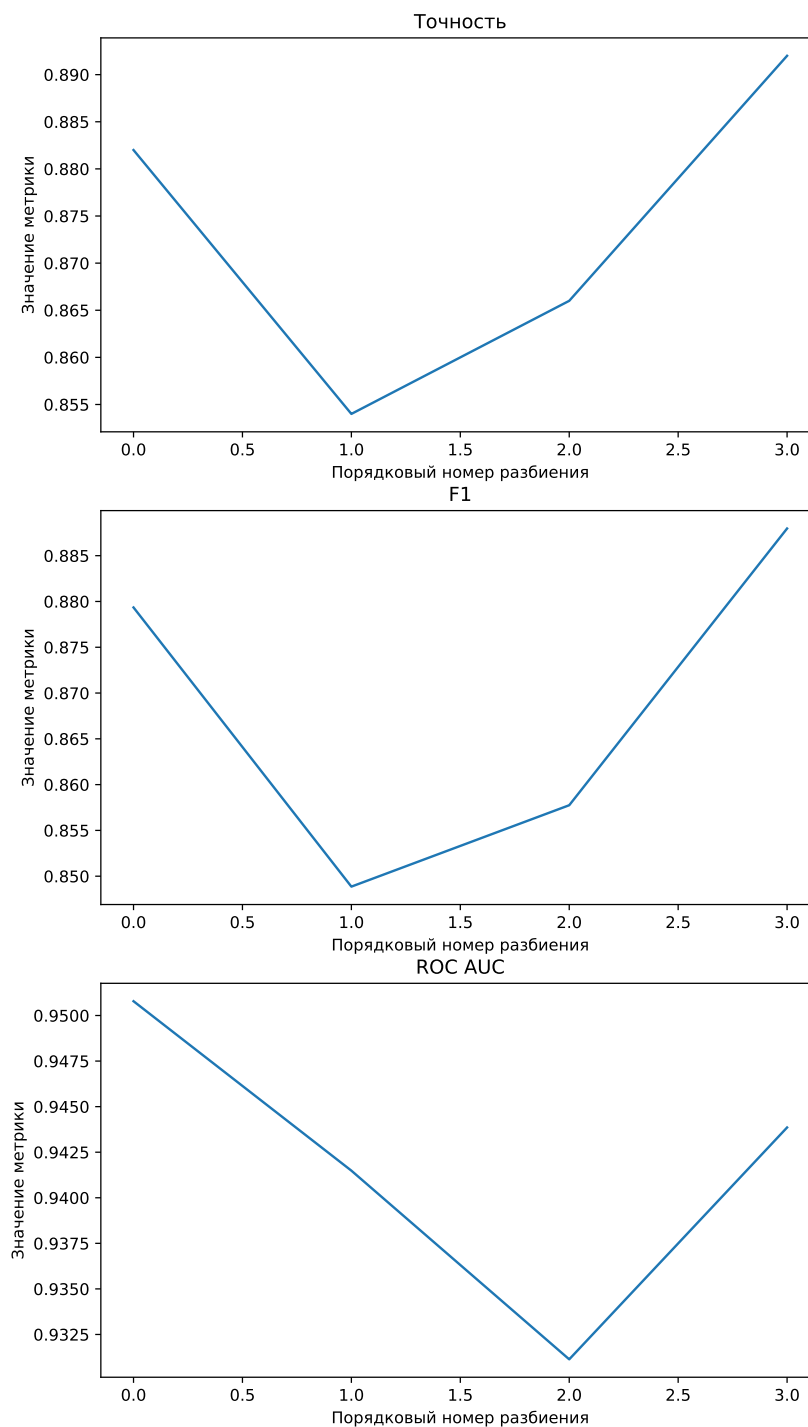


Рисунок 4.20 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием логистической регрессии (метод векторизации – BERT).

#### 4.2.6 Перцептрон

Параметры модели при применении метода векторизации “мешок слов”:

- скорость обучения – 0.0001;
- максимальное количество итераций при обучении – 30;
- штраф – сумма квадратов весов модели, умноженных на гиперпараметр регуляризации.

На рисунке 4.21 представлены матрицы ошибок, полученные с использованием перцептрона, метод векторизации – “мешок слов”.

На рисунке 4.22 представлены оценки классификатора, полученные с использованием перцептрона, метод векторизации – “мешок слов”.

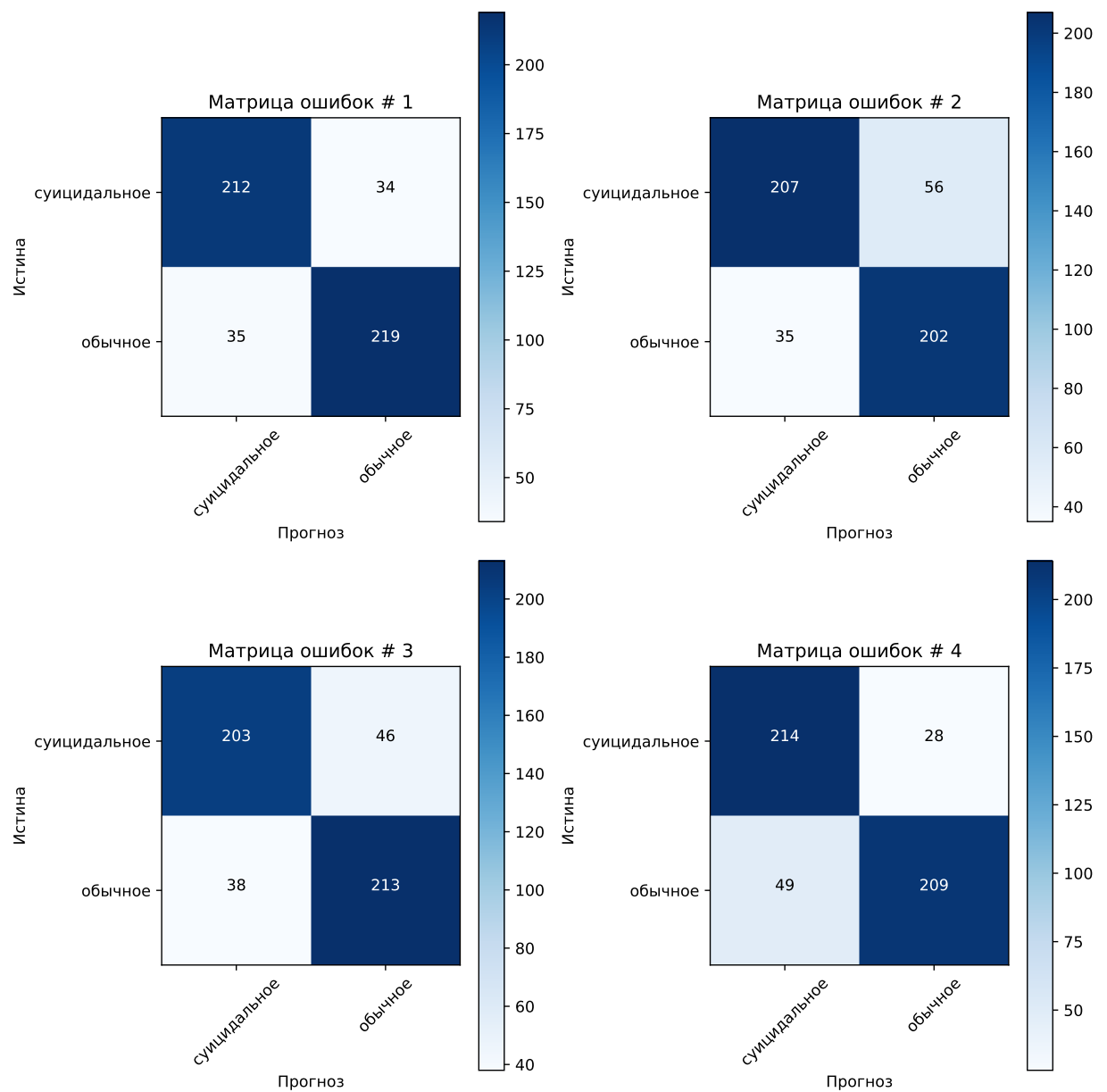


Рисунок 4.21 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием перцептрона (метод векторизации – “мешок слов”).

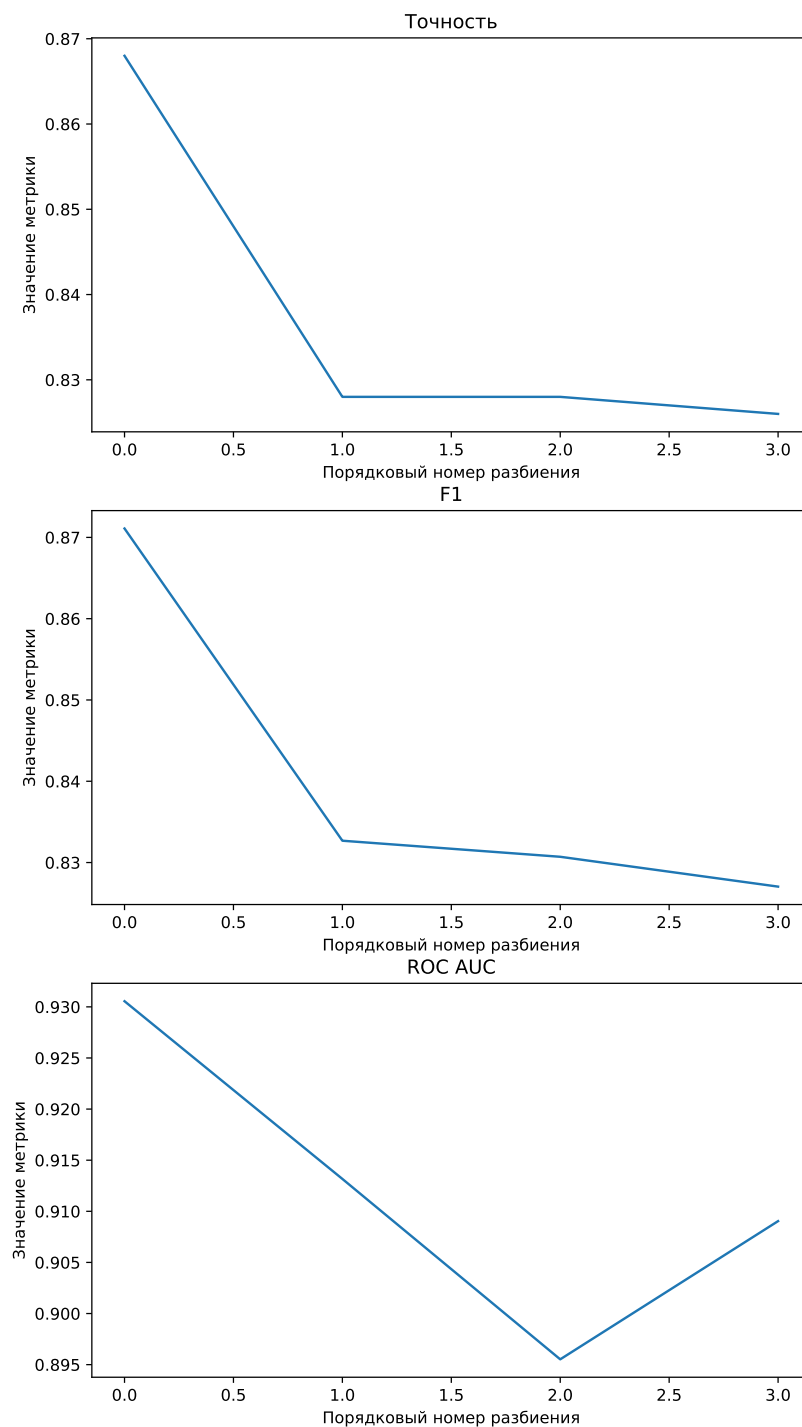


Рисунок 4.22 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием перцептрона (метод векторизации – “мешок слов”).



Параметры модели при применении векторизации BERT:

- скорость обучения – 0.0001;
- максимальное количество итераций при обучении – 30;
- штраф – абсолютное значение коэффициентов модели.

На рисунке 4.23 представлены матрицы ошибок, полученные с использованием перцептрона, метод векторизации – BERT.

На рисунке 4.24 представлены оценки классификатора, полученные с использованием перцептрона, метод векторизации – BERT.

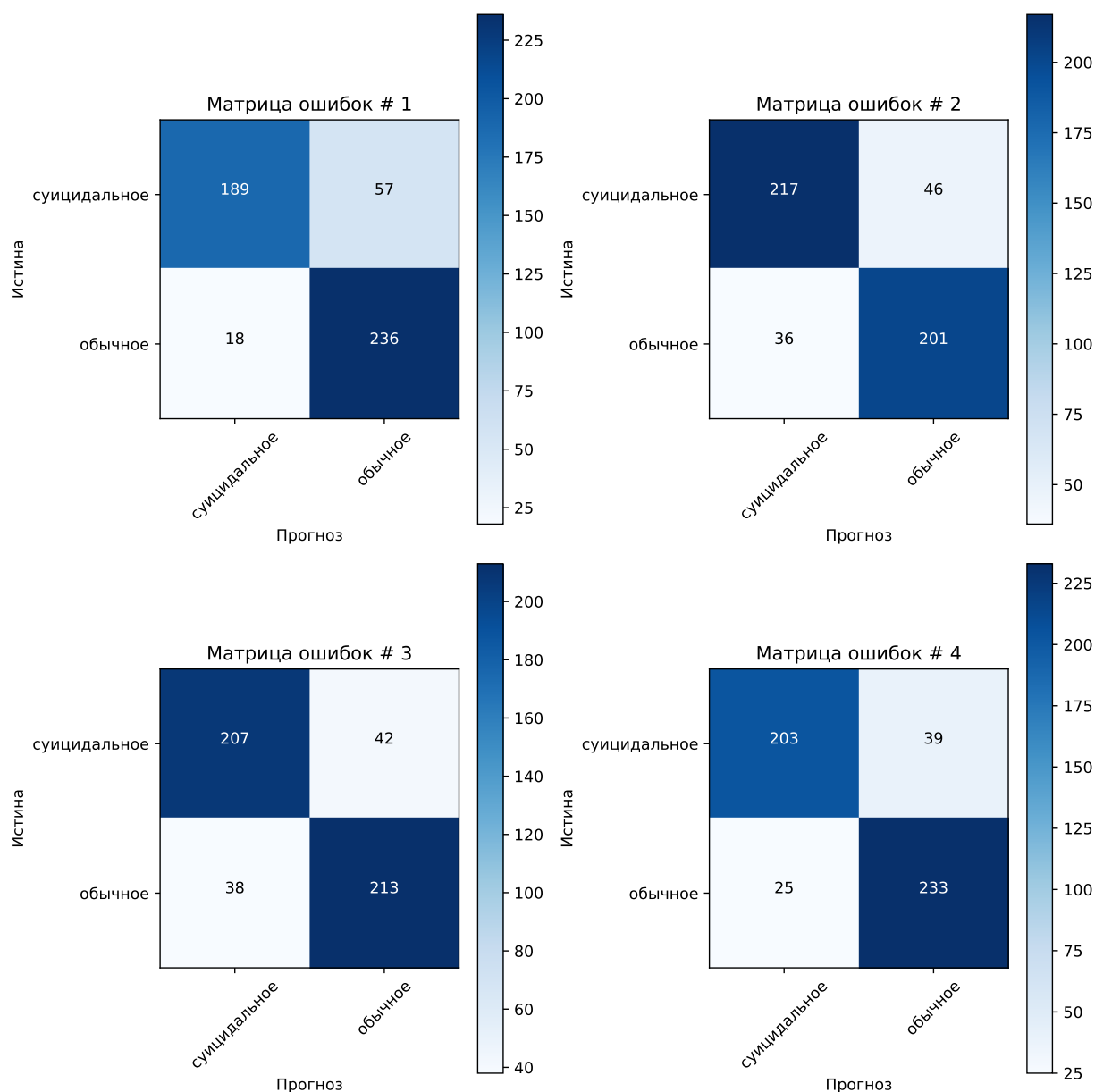


Рисунок 4.23 – Матрицы ошибок в зависимости от номера разбиения, полученные с использованием перцептрона (метод векторизации – BERT).

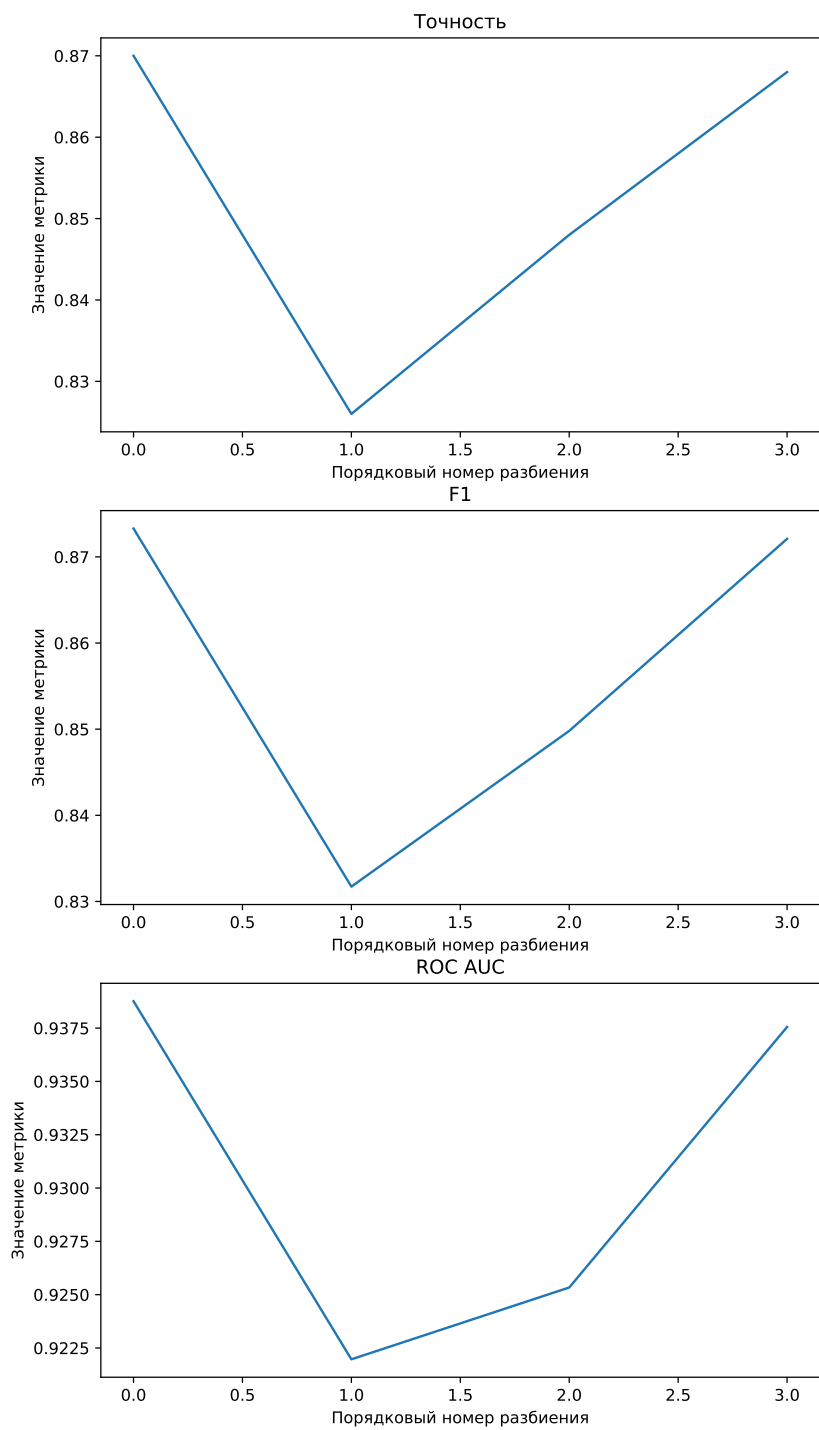


Рисунок 4.24 – Оценки классификатора в зависимости от номера разбиения, полученные с использованием перцептрона (метод векторизации – BERT).

## Вывод

По представленным матрицам ошибок можно увидеть, что все методы, кроме случайного леса (BERT), метода опорных векторов (оба варианта векторизации) и перцептрона (оба варианта), в большинстве своем в 2 раза чаще определяют обычное сообщение как суицидальное. Метод опорных векторов имеет более распределенные матрицы ошибок. Самые частые ошибки в определении обычных сообщений совершал метод К-ближайших соседей. Перцептрон на каждом из разбиений мог либо хуже распознавать суицидальные, либо хуже распознавать обычные сообщения.

В таблицах 2 и 3 приведены результаты проведенного исследования, в которых для каждого алгоритма выделены лучшие показатели каждой метрики.

Из таблиц видно, что лучшим средним показателем метрик точности, F1-меры и ROC-AUC обладает метод случайного леса с использованием BERT-векторизации. Его метрика точности достигла показателя 0.888, F1-мера – 0.886, а ROC-AUC – 0.949. При этом на втором месте располагается тот же метод, но с использованием векторизации “Мешок слов”, относительно первого метода его точность уступает на  $\approx 0.4\%$ , F1-мера – на  $\approx 1.4\%$ , а ROC-AUC – на  $\approx 0.2\%$ . На третьем месте располагается логистическая регрессия с использованием BERT-векторизации, его точность ниже на  $\approx 1.6\%$ , F1-мера – на  $\approx 1.9\%$ , а ROC-AUC – на  $\approx 0.7\%$ . Таким образом, в качестве используемой модели в задаче распознавания паттернов суицидального поведения человека по суицидальным сообщениям лучше всего воспользоваться именно методом случайного леса с указанными в исследовании параметрами.

Стоит отметить, что выбранный метод также в  $\approx 1.5$  раза чаще ошибочно интерпретирует обычные сообщения как суицидальные, чем суицидальные как обычные. Данный факт не относится к проблеме модели, которая может помешать работе системы в силу того, что распознавание суицидальных сообщений для нее играет первостепенную роль.

Кроме того, с использованием полученной информации можно предложить использовать ансамбль задействованных в исследовании моделей, так как результаты их работы имеют достаточно высокие показатели F1-меры в сбалансированном наборе данных.

Таблица 2 – Результаты исследования

Алгоритм	Векторизация	Точность	F1-мера	ROC-AUC
Градиентный бустинг	“Мешок слов”	0.858	0.849	0.926
		0.836	0.830	0.909
		0.854	0.841	0.902
		0.864	0.854	0.926
	<i>среднее</i>	<i>0.853</i>	<i>0.844</i>	<i>0.916</i>
	BERT	0.868	0.849	0.934
		0.852	0.840	0.918
		0.870	0.852	0.924
		0.890	0.876	0.939
	<i>среднее</i>	<i>0.870</i>	<i>0.854</i>	<i>0.929</i>
Случайный лес	“Мешок слов”	0.888	0.876	0.951
		0.876	0.858	0.942
		0.884	0.881	0.937
		0.892	0.877	0.959
	<i>среднее</i>	<i>0.885</i>	<i>0.873</i>	<i>0.947</i>
	BERT	0.886	0.884	0.954
		0.888	0.884	0.946
		0.878	0.872	0.936
		0.900	0.902	0.960
	<i>среднее</i>	<b><i>0.888</i></b>	<b><i>0.886</i></b>	<b><i>0.949</i></b>
Метод опорных векторов	“Мешок слов”	0.850	0.851	0.924
		0.848	0.847	0.920
		0.824	0.819	0.899
		0.864	0.862	0.918
	<i>среднее</i>	<i>0.847</i>	<i>0.845</i>	<i>0.915</i>
	BERT	0.864	0.865	0.930
		0.850	0.850	0.920
		0.848	0.844	0.919
		0.886	0.884	0.928
	<i>среднее</i>	<i>0.862</i>	<i>0.861</i>	<i>0.924</i>

Таблица 3 – Результаты исследования

Алгоритм	Векторизация	Точность	F1-мера	ROC-AUC
Метод К-ближайших соседей	“Мешок слов”	0.742	0.692	0.854
		0.750	0.691	0.844
		0.738	0.678	0.845
		0.778	0.754	0.874
	<i>среднее</i>	<i>0.752</i>	<i>0.704</i>	<i>0.854</i>
	BERT	0.772	0.749	0.868
		0.762	0.740	0.831
		0.720	0.696	0.816
		0.778	0.761	0.853
	<i>среднее</i>	<i>0.758</i>	<i>0.737</i>	<i>0.842</i>
Логистическая регрессия	“Мешок слов”	0.864	0.858	0.947
		0.860	0.853	0.939
		0.862	0.851	0.926
		0.880	0.872	0.936
	<i>среднее</i>	<i>0.867</i>	<i>0.859</i>	<i>0.937</i>
	BERT	0.882	0.879	0.951
		0.854	0.849	0.941
		0.866	0.858	0.931
		0.892	0.888	0.944
	<i>среднее</i>	<i>0.874</i>	<i>0.869</i>	<i>0.942</i>
Перцептрон	“Мешок слов”	0.868	0.871	0.931
		0.828	0.833	0.913
		0.828	0.831	0.896
		0.826	0.827	0.909
	<i>среднее</i>	<i>0.838</i>	<i>0.841</i>	<i>0.912</i>
	BERT	0.870	0.873	0.939
		0.826	0.832	0.922
		0.848	0.850	0.925
		0.868	0.872	0.938
	<i>среднее</i>	<i>0.853</i>	<i>0.857</i>	<i>0.931</i>

## ЗАКЛЮЧЕНИЕ

Были рассмотрены термины предметной области, включающие в себя понятия самоубийства и суицидального поведения. Представлен интегративный теоретический подход в суицидологии, описаны суицидальное поведение, внешние проявления суицидального поведения, пресуицид, парасуицид и постсуицид. Приведена информация об определении истинности суицидальных намерений.

Описаны факторы повышенного суицидального риска и мероприятия и методики предотвращения самоубийств, включающие в себя суицидологическую диагностику, кризисную терапию и Телефон Доверия. Представлена статистика совершения самоубийств.

С использованием классификации признаков паттернов суицидального поведения человека описаны форматы хранения проявления поведения человека. Были выделены аудиальные, текстовые, пространственно-временные, визуальные, физиологические и биологические признаки.

Был описан метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям, а также формат и метод сбора задействованных в нем данных. В качестве средства сбора данных использован бот в мессенджере Telegram. Рассмотрены доступные средства реализации ботов в выбранном мессенджере.

Был определен перечень задействованных методов машинного обучения, который включил в себя: градиентный бустинг, метод случайного леса, метод опорных векторов, метод К-ближайших соседей, логистическую регрессию и перцептрон. В качестве методов векторизации выбраны: алгоритм “мешок слов” и языковая модель BERT.

Была приведена диаграмма вариантов использования. Для системы было определено три действующих лица: пользователь, рекомендатор и анализатор. Приведена IDEF0 диаграмма, декомпозирована главная задача метода – распознавание суицидального сообщения. Диаграмма “сущность-связь” в нотации Чена позволила на абстрактном уровне описать систему распознавания.

Был разработан метод распознавания суицидальных паттернов поведения человека по текстовым сообщениям, который включил в себя хранение

и анализ сообщений пользователей. Для определения, является ли сообщение суицидальным, используется модель машинного обучения. В качестве обучающей выборки используется дополненный датасет размеченных несуицидальных сообщений из открытого доступа.

Разработанный метод был реализован. Представлены интерфейсы средства сбора данных и средства распознавания суицидальных паттернов поведения человека по текстовым сообщениям.

Представленные диаграммы тональности сообщений показали, что практически треть суицидальных сообщений автоматизированное средство оценки тональности распознает как сообщения с отрицательной окраской. Среди несуицидальных сообщений преобладают тексты с отрицательной окраской, при этом нейтральных сообщений – четверть из всех.

Визуализированные облака слов подтвердили гипотезу, что выбранные классы суицидальных и несуицидальных сообщений разделимы и отличны частотой некоторых слов. Отмечено, что слово “хотеть” встречается в суицидальных сообщениях в  $\approx 7.83$  раза чаще, чем в несуицидальных, а слово “человек” – в  $\approx 7.33$  раза чаще. Таким образом, суицидальные сообщения являются менее “разнообразными” и фиксирующимися на определенном словарном множестве.

Было проведено исследование, включившее в себя построение матриц ошибок и определение метрик точности, F1-меры и ROC-AUC для каждого рассматриваемого алгоритма машинного обучения. Для определения гиперпараметров каждой модели применялся метод поиска по сетке с опорой на значение F1-меры. Разбиение данных на выборки производилось на 4 части, 1 из которых используется в качестве тестовой. Лучшее среднее значение всех метрик показал метод случайного леса с использованием BERT-векторизации. Его точность достигла показателя 0.888, F1-мера – 0.887, а ROC-AUC – 0.948. При этом на втором месте располагается тот же метод, но с использованием векторизации “Мешок слов”, относительно первого метода его точность уступила на  $\approx 1.1\%$ , F1-мера – на  $\approx 1.4\%$ , а ROC-AUC – на  $\approx 0.1\%$ . На третьем месте располагается логистическая регрессия с использованием BERT-векторизации, его точность ниже на  $\approx 1.6\%$ , F1-мера – на  $\approx 2.1\%$ , а ROC-AUC – на  $\approx 0.6\%$ . Таким образом, в качестве используемой модели в задаче распознавания суицидального поведения человека по

текстовым сообщениям рекомендуется использовать метод случайного леса с указанными в исследовании параметрами. Стоит отметить, что данный метод также на  $\approx 67\%$  чаще ошибочно интерпретирует обычные сообщения как суицидальные, чем суицидальные как обычные. Данный факт не относится к проблеме модели, которая может помешать работе системы в силу того, что распознавание суицидальных сообщений для нее играет первостепенную роль.



# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

## Список литературы

1. Suicide: facts and figures globally [Электронный ресурс]. Режим доступа: <https://www.who.int/publications/i/item/WHO-MSD-UCN-MHE-22.03> (дата обращения 08.05.2023).
2. В.А. Аверин. Интегративный подход и принципы психологии развития человека // Психология человека в образовании. 2019. № 3.
3. Cambridge Dictionary [Электронный ресурс]. Режим доступа: <https://dictionary.cambridge.org/dictionary/english/suicide> (дата обращения 08.05.2023).
4. WHO: Suicide [Электронный ресурс]. Режим доступа: <https://www.who.int/news-room/fact-sheets/detail/suicide#:~:text=Key%20facts,suicide%20in%20the%20general%20population.> (дата обращения 08.05.2023).
5. Брябрина Т.В. Гиберт А.И. Штрахова А.В. Опыт контент-анализа суицидальных высказываний в сети Интернет лиц с различным уровнем суицидальной активности // Психология. Психофизиология. 2016. № 3.
6. Г.В. Старшенбаум. Клиническая психология. Суицидология и кризисная терапия. М.: Когито-Центр, 2018. с. 376.
7. Е.В. Гришина. Психологические характеристики образовательной среды и антивитальные переживания подростков // Народное образование. 2016. № 4.
8. В.В. Касьянов. Суицидальное поведение. М.: Юрайт, 2019. с. 330.
9. Амбрумова А.Г. Бородин С.В. Михлин А.С. Предупреждение самоубийств. М.: Издательство Академии МВД СССР, 1980. с. 164.
10. В.А. Тихоненко. Классификация суицидальных проявлений // Актуальные проблемы суицидологии. Труды Московского НИИ психиатрии. 1978.

11. L. Rappoport. The state of crisis: some theoretical considerations // Social Service Review. 1962. Т. 36, № 3.
12. М.Я. Соловейчик. Кризис и основы кризисной интервенции // Мастерство психологического консультирования. 2002.
13. Centers for Disease Control and Prevention: Suicide Prevention [Электронный ресурс]. Режим доступа: <https://www.cdc.gov/suicide> (дата обращения 08.09.2023).
14. Г.В. Порядина. Стресс и патология: учеб. пособие. М.: РГМУ, 2009. с. 23.
15. Овсянников А.А. Грызлов И.Н. Голубинский Е.Ю. Подход к автоматической классификации коротких сообщений на основе модифицированного метода Байеса // Экономика. Информатика. 2014. № 8-1.
16. Катермина Т.С. Тагиров Т.М. Элементы искусственного интеллекта в решении задач анализа текстов // Computational nanotechnology. 2022. № 2.
17. Riezler S. Hagmann M. Validity, reliability, and significance: Empirical methods for NLP and data science // Synthesis Lectures on Human Language Technologies. 2022. № 14.
18. Zhang T. Schoene A.M. Ji S. Natural language processing applied to mental illness detection: A narrative review // Npj Digital Medicine. 2022. № 5.
19. Khurshid S. Reeder C. Harrington L. Cohort design and natural language processing to reduce bias in electronic health records research // Npj Digital Medicine. 2022. Т. 5, № 47.
20. Orea-Giner A. Fuentes-Moraleda L. Does the implementation of robots in hotels influence the overall tripadvisor rating? A text mining analysis from the Industry 5.0 approach // Tourism Management. 2022. № 93.
21. Ledro C. Nosella A. Vinelli A. Artificial intelligence in customer relationship management: Literature review and future research directions // Journal of Business and Industrial Marketing. 2022. № 37.

22. Nijhawan T. Attigeri G. Ananthakrishna T. Stress detection using natural language processing and machine learning over social interactions // Journal of Big Data. 2022. № 9.
23. Буянов И. Соченков И. Датасет для задачи распознавания пресуицидальных сигналов в тексте и его анализ // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2022”. 2022.
24. Gradient Boosting and XGBoost [Электронный ресурс]. Режим доступа: <https://medium.com/@gabrieltseng/gradient-boosting-and-xgboost-c306c1bcfaf5> (дата обращения 10.05.2023).
25. How Facebook AI Helps Suicide Prevention [Электронный ресурс]. Режим доступа: <https://about.fb.com/news/2018/09/inside-feed-suicide-prevention-and-ai/> (дата обращения 14.04.2024).
26. Mendeley Data: Dataset for presuicidal signal detection. Режим доступа: <https://data.mendeley.com/datasets/86v3z38dc7/1> (дата обращения 11.09.2023).
27. Федеральный закон от 27.07.2006 г. № 152-ФЗ “О персональных данных” // СПС “КонсультантПлюс”.
28. Python-telegram-bot official page [Электронный ресурс]. Режим доступа: <https://python-telegram-bot.org/> (дата обращения 10.12.2023).
29. Telebot documentation [Электронный ресурс]. Режим доступа: <https://pytba.readthedocs.io/en/latest/> (дата обращения 10.12.2023).
30. Официальная страница модуля node-telegram-bot-api [Электронный ресурс]. Режим доступа: <https://www.npmjs.com/package/node-telegram-bot-api> (дата обращения 10.12.2023).

31. Репозиторий проекта kotlin-telegram-bot [Электронный ресурс]. Режим доступа: <https://github.com/kotlin-telegram-bot/kotlin-telegram-bot> (дата обращения 10.12.2023).
32. Python official page [Электронный ресурс]. Режим доступа: <https://www.python.org/> (дата обращения 10.05.2023).
33. mdn web docs: JavaScript [Электронный ресурс]. Режим доступа: <https://developer.mozilla.org/en-US/docs/Web/JavaScript> (дата обращения 10.12.2023).
34. Node.js documentation [Электронный ресурс]. Режим доступа: <https://nodejs.org/docs/latest/api/> (дата обращения 10.12.2023).
35. Kotlin language specification [Электронный ресурс]. Режим доступа: <https://kotlinlang.org/spec/introduction.html> (дата обращения 09.10.2020).
36. Java Virtual Machine SE8 Specification [Электронный ресурс]. Режим доступа: <https://docs.oracle.com/javase/specs/jvms/se8/html/> (дата обращения 08.09.2023).
37. Спивак А.И. Лапшин С.В. Лебедев И.С. Классификация коротких сообщений с использованием векторизации на основе ELMo // Известия ТулГУ. Технические науки. 2019. № 10.
38. Zhang Y. Jin R. Zhou Z.H. Understanding bag-of-words model: a statistical framework // International Journal of Machine Learning and Cybernetics. 2010. Т. 1.
39. K.S. Jones. A statistical interpretation of term specificity and its application in retrieval // Journal of Documentation. 2004. Т. 60, № 1.
40. Tomas Mikolov Kai Ceen. Efficient Estimation of Word Representations in Vector Space. 2013.
41. Peters M.E. Neumann M. Iyyer M. Deep contextualized word representations // Proceedings of the 2018 Conference of the North American

Chapter of the Association of Computational Linguistics: Human Language Technologies. 2018. T. 1.

42. Cornell University arxiv [Электронный ресурс]. Режим доступа: <https://arxiv.org/abs/1810.04805v2> (дата обращения 10.05.2023).
43. PostgreSQL: The World's Most Advanced Open Source Relational Database (official page) [Электронный ресурс]. Режим доступа: <https://www.postgresql.org/> (дата обращения 24.03.2024).
44. Pandas official page [Электронный ресурс]. Режим доступа: <https://pandas.pydata.org/> (дата обращения 10.05.2023).
45. Numpy official page [Электронный ресурс]. Режим доступа: <https://numpy.org/> (дата обращения 10.05.2023).
46. Matplotlib official page [Электронный ресурс]. Режим доступа: <https://matplotlib.org/> (дата обращения 10.05.2023).
47. Scikit-learn official page [Электронный ресурс]. Режим доступа: <https://scikit-learn.org/stable/> (дата обращения 10.05.2023).
48. NLTK Documentation [Электронный ресурс]. Режим доступа: <https://www.nltk.org/> (дата обращения 24.03.2024).
49. Морфологический анализатор rymorphy2, документация [Электронный ресурс]. Режим доступа: <https://pymorphy2.readthedocs.io/en/stable/> (дата обращения 24.03.2024).
50. Репозиторий проекта Dostoevsky [Электронный ресурс]. Режим доступа: <https://github.com/bureaucratic-labs/dostoevsky> (дата обращения 06.04.2024).