

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение
высшего образования
«ТЮМЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ НАУК
Кафедра программной и системной инженерии

РЕКОМЕНДОВАНО К ЗАЩИТЕ В
ГЭК

Заведующий кафедрой
д.т.н., профессор

_____ А.Г. Ивашко
_____ 2023 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
магистра

**РАЗРАБОТКА ПОДСИСТЕМЫ ПРОГНОЗИРОВАНИЯ УХУДШЕНИЯ
СОСТОЯНИЯ ЗДОРОВЬЯ ПАЦИЕНТОВ С СЕРДЕЧНО-СОСУДИСТЫМИ
ЗАБОЛЕВАНИЯМИ**

09.04.03 Прикладная информатика

Магистерская программа «Информационные системы анализа данных»

Выполнил работу
студент 2 курса
очной формы обучения



Романов
Данил
Александрович

(подпись)

Руководитель
доцент, к. т. н.



Цыганова
Мария
Сергеевна

(подпись)

Рецензент
к. ф.-м. н., доцент



Семихин
Дмитрий
Витальевич

(подпись)

Тюмень, 2023

ОГЛАВЛЕНИЕ

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ.....	4
ВВЕДЕНИЕ	5
1. ОПИСАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ.....	7
1.1 Проблематика.....	9
1.2 Цели	9
1.3 Задачи.....	10
2. РАЗРАБОТКА МОДЕЛИ ПРОГНОЗИРОВАНИЯ.....	11
2.1 Формирование набора данных	12
2.1.1 Описание объектов набора данных	12
2.1.2 Описание признаков.....	13
2.1.3 Реализация экспорта набора данных	18
2.1.4 Выгрузка.....	23
2.2 Сравнительный анализ алгоритмов машинного обучения	26
2.2.1 Логистическая регрессия	26
2.2.2 К-ближайшие соседи.....	27
2.2.3 Случайный лес	28
2.2.4 Одноклассовая классификация SVM.....	28
2.2.5 Градиентный бустинг.....	29
2.2.6 Вывод.....	30
2.3 Реализация предварительной обработки данных	30
2.3.1 Используемые технические средства и программы.....	30
2.3.2 Процесс предобработки данных	31

2.4	Реализация модели градиентного бустинга	37
2.4.1	Используемые технические средства и программы.....	37
2.4.2	Подбор гиперпараметров модели	38
2.4.3	Оценки качества обученной модели прогнозирования	38
2.5	Вывод.....	41
3.	РЕАЛИЗАЦИЯ ВЕБ-СЕРВИСА.....	43
3.1	Используемые технические средства и программы.....	43
3.2	Разработка	43
3.3	Вывод.....	47
4	РАЗРАБОТКА ПОДСИСТЕМЫ МИС «1С:Медицина.Регион».....	48
4.1	Проектирование	48
4.2	Реализация.....	54
4.3	Вывод.....	69
5	Заключение.....	70
	Список литературы.....	71
	ПРИЛОЖЕНИЕ 1	73

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

МО – медицинская организация

МУ – медицинское учреждение

ССЗ – сердечно-сосудистые заболевания

МИС – медицинская информационная система

ИБС – ишемическая болезнь сердца

ХСН – хроническая сердечная недостаточность

ОКБ – острый коронарный синдром

ВВЕДЕНИЕ

Информационные процессы присутствуют во всех областях медицины и здравоохранения. Современные информационные технологии все больше используются в медицинских организациях (МО). В некоторых медицинских исследованиях невозможно обойтись без компьютера и специального медицинского оборудования. Благодаря этому, практически в каждой МО для автоматизации ее деятельности используется медицинская информационная система (МИС).

Сейчас на рынке представлено больше сотни различных систем для медицинских учреждений (МУ). На первой странице поискового запроса отображаются: Medesk, Mework, Interin Medods, «1С:Медицина.Больница». Выбор МИС зависит от типа МО, количества сотрудников, состава отделений организации, особенностей бизнес-процессов и др.

На текущий момент, практически во всех государственных МУ Тюмени и Тюменской области активно используются и продвигаются МИС «1С:Медицина.Регион».

В настоящее время большую популярность приобрели технологии искусственного интеллекта, включающие алгоритмы машинного обучения, искусственные нейронные сети и глубокое обучение.

Отдельного внимания заслуживает медицина и применение технологий машинного обучения в данной сфере. Трудно не согласиться, что социальное и экономическое благополучие общества в значительной степени зависит от уровня развития системы здравоохранения в целом, и качества медицинского обслуживания населения, в частности. На текущий момент, существует множество решений в медицине, использующих анализ данных и машинное обучение, применяющихся для решения задач прогнозирования и классификации. Посредством общения с пользователями системы – врачом персоналом – открываются новые области и идеи для применения интеллектуального анализа данных. Информационные системы, основанные на машинных алгоритмах, позволяют более точно диагностировать

заболевания, выбрать оптимальную стратегию лечения, просчитать риски, варианты развития болезни или ее возможный исход. [1]

Для успешной подготовки и защиты выпускной квалификационной работы обучающимся использовались средства и методы физической культуры и спорта с целью поддержания должного уровня физической подготовленности, обеспечивающую высокую умственную и физической работоспособность. В режим рабочего дня включались различные формы организации занятий физической культурой (физкультпаузы, физкультминутки, занятия избранным видом спорта) с целью профилактики утомления, появления хронических заболеваний и нормализации деятельности различных систем организма.

В рамках подготовки к защите выпускной квалификационной работы автором созданы и поддерживались безопасные условия жизнедеятельности, учитывающие возможность возникновения чрезвычайных ситуаций.

1. ОПИСАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ

На сегодняшний день, одним из направлений, которым отводится больше всего внимания, являются заболевания сердечно-сосудистой системы (ССЗ). Согласно актуальному проведенному анализу в рамках «Федерального проекта по борьбе с сердечно-сосудистым заболеваниям» по Тюменской области, динамика смертности от ССЗ с 2020 года по 2021 составила -1.9% (511.7 на 100 тысяч человек), а в процентном соотношении причин смерти данная группа заболеваний занимает первое место и составляет 37.1%. Распределение умерших от болезней системы кровообращения по месту смерти следующее: 40,4% - в стационаре, 59,6% - вне стационара. [3]

Из-за большого риска ухудшения состояния здоровья пациентов и развития осложнений больных ССЗ, большую роль для снижения смертности населения выполняют оперативная диагностика и профилактика. Для этого разрабатываются различные меры по повышению качества оказания медицинской помощи у пациентов ключевых групп ССЗ, определяющие основной вклад в заболеваемость и смертность. [2]

Сердечно-сосудистые заболевания представляют собой группу болезней сердца и кровеносных сосудов, в которую входят [4]:

- ишемическая болезнь сердца – болезнь кровеносных сосудов, снабжающих кровью сердечную мышцу;
- болезнь сосудов головного мозга – болезнь кровеносных сосудов, снабжающих кровью мозг;
- болезнь периферических артерий – болезнь кровеносных сосудов, снабжающих кровью руки и ноги;
- ревмокардит – поражение сердечной мышцы и сердечных клапанов в результате ревматической атаки, вызываемой стрептококковыми бактериями;
- врожденный порок сердца – существующие с рождения деформации строения сердца;
- тромбоз глубоких вен и эмболия легких – образование в ножных венах сгустков крови, которые могут смещаться и двигаться к сердцу и легким;

- другие.

В структуре смертности от болезней системы кровообращения в 2021 году на первом месте – ишемические болезни сердца. Доля данной причины составляет 69,5% (число умерших на 100 тыс. населения – 355,7) от общего числа умерших от болезней системы кровообращения. На втором месте – цереброваскулярные заболевания. Доля данной причины составляет 15,3% (число умерших на 100 тыс. населения – 78,2) от общего числа умерших от болезней системы кровообращения. Третье место в структуре смертности от болезней системы кровообращения занимает кардиомиопатия, и доля данной причины занимает 9,3% (число умерших на 100 тыс. населения – 46,8). [3]

По данным на 2021 год по классу болезней ССЗ и пациентам в возрасте 18 лет и старше:

- под диспансерное наблюдение было взято 373 126 случаев, из которых 83 797 – по поводу ИБС;
- в медицинских организациях Тюменской области оказано 43 042 случая медицинской помощи в стационарных условиях;
- было выписано 228 963 пациента, из которых 10515 умерло.

Высокая роль и значимость повышения качества оказываемой медицинской помощи больным ССЗ подтверждается национальным проектом «Здравоохранение», разработанным в целях выполнения указа Президента РФ №204 от 07.05.2018 «О национальных целях и стратегических задачах развития Российской Федерации на период до 2024 года.

Некоторыми из проектов, разработанных на основании указа, являются вертикально-интегрированная медицинская информационная система (ВИМИС) и «Борьба с сердечно-сосудистыми заболеваниями», одной из задач которых является мониторинг и контроль своевременного оказания медицинской помощи пациентам с ССЗ [1].

Однако, перед оказанием амбулаторной или стационарной медицинской помощи или назначения профилактических действий, пользователи МИС «1С:Медицина.Регион», а именно лечащие врачи, должны диагностировать

само заболевание, проанализировав как различные показатели здоровья, так и общее физическое состояние пациента, которые в виду особенностей ССЗ, изменяются крайне быстро.

1.1 Проблематика

Для проведения профилактики или диагностирования заболевания первоначально лечащим врачом амбулатории назначается лабораторный анализ определенных видов показателей здоровья, который в последствии выполняется пациентом.

Контроль и назначение лечения, а также решение о плановой госпитализации пациента проводятся лечащим врачом амбулатории после получения результатов лабораторных исследований и последующего повторного приема пациента, между которыми может проходить достаточно большое количество времени.

Не своевременное оказание медицинской помощи как амбулаторной, так и стационарной, влечет за собой огромный риск ухудшения состояния здоровья пациента или развития осложнений, которые в большинстве случаев приводят к экстренной госпитализации, создавая дополнительную нагрузку на СМП и стационар.

В данном случае, наиболее «опасными» диагнозами ССЗ, приводящим к неблагоприятным последствиям, являются: ишемическая болезнь сердца (ИБС), острый коронарный синдром (ОКС) и хроническая сердечная недостаточность (ХСН).

1.2 Цели

При исследовании проблематики, были сформированы следующие цели:

- сократить время на оказание медицинской помощи больным ССЗ, разработав подсистему прогнозирования ухудшения состояния здоровья пациентов с ССЗ;
- сократить количество экстренных госпитализаций, вызовов неотложной медицинской помощи.

1.3 Задачи

На основании поставленной цели, были сформулированы следующие задачи:

- изучить и отобрать перечень видов показателей здоровья, значимых при принятии решения о госпитализации пациента с предварительным диагнозом ИБС, ХСН и ОКС;
- определить случаи, описывающие наличие и отсутствие риска ухудшения состояния пациента;
- сформировать первичный набор данных из показателей здоровья;
- на основе анализа имеющихся данных спроектировать и реализовать механизм предварительной обработки, проанализировать данные, проверить наличие выбросов и пропусков, реализовать процедуры предварительной обработки данных, выбрав методы/стратегии для каждого поля из полученного списка признаков;
- провести сравнительный анализ имеющихся алгоритмов машинного обучения для текущей задачи. Выбрать, создать и обучить модель машинного обучения;
- протестировать и проанализировать результат работы модели, в случае необходимости подобрать гиперпараметры так, чтобы добиться лучших результатов, получить итоговые оценки;
- рассмотреть и проанализировать итоговые метрики работы модели, определить корректность использования ее в рабочей системе;
- разработать веб-сервис для взаимодействия с моделью;
- спроектировать и разработать подсистему прогнозирования ухудшения здоровья пациента со стороны МИС «1С:Медицина.Регион»;
- протестировать общую работу подсистемы и HTTP - сервиса.

2. РАЗРАБОТКА МОДЕЛИ ПРОГНОЗИРОВАНИЯ

Задача прогнозирования представляет предсказание будущих событий. Относительно текущих сформулированных целей, прогнозирование заключается в предсказании необходимости плановой госпитализации пациента, другими словами – прогнозирование риска ухудшения состояния здоровья.

Всего в задаче прогнозирования будут рассмотрены три группы ССЗ:

- ишемическая болезнь сердца;
- острый коронарный синдром;
- хроническая сердечная недостаточность.

Ишемическая болезнь сердца (далее ИБС) - болезнь, при которой у человека нарушается кровоснабжения миокарда из-за патологий коронарных артерий. Данное заболевание является самой распространённой причиной скоропостижной смерти людей трудоспособного возраста [2] и занимает ведущее место по своей медико-социальной значимости и влиянию на общую заболеваемость, нарушение трудоспособности, инвалидизацию во всех странах мира.

Наибольший риск для пациента при наличии ИБС представляет развитие острого коронарного синдрома (далее ОКС). ОКС является самым опасным вариантом клинического течения ИБС. Опасность ситуации состоит в том, что ОКС, в отличие от хронической ИБС, характеризуется быстрым (часы), а иногда стремительным (минуты) течением болезни, высоким риском неблагоприятных исходов (внезапная коронарная смерть, инфаркт миокарда) и требует безотлагательных мер по спасению жизни больных. Термин «острый коронарный синдром» объединяет такие клинические состояния как нестабильная стенокардия и инфаркт миокарда. [3]

Другим серьезным заболеванием является «Хроническая сердечная недостаточность» (далее ХСН) - одно из самых тяжелых заболеваний сердечно-сосудистой системы, имеющих плохой прогноз, как для жизни, так

и для трудовой деятельности. После постановки первоначального диагноза пациенты с сердечной недостаточностью госпитализируются в среднем один раз в год. Крайне неблагоприятный прогноз и низкая выживаемость больных, которые сопоставимы или даже хуже аналогичных показателей при онкологических заболеваниях составляют значительную нагрузку на систему здравоохранения. [4]

2.1 Формирование набора данных

Каждое из вышерассмотренных заболеваний представляет из себя группу следующих диагнозов МКБ-10:

- ХСН – I50.0, I50.1, I50.9;
- ИБС – I20.1, I20.8, I20.9, I25.0, I25.1, I25.2, I25.3, I25.4, I25.5, I25.6, I25.8, I25.9;
- ОКС – I20.0, I21.0, I21.1, I21.2, I21.3, I21.4, I21.9, I22.0, I22.1, I22.8, I22.9, I24.0, I24.8, I24.9.

2.1.1 Описание объектов набора данных

В качестве объектов набора данных для задачи прогнозирования были отобраны следующие виды случаев направлений в стационар с вышеуказанными диагнозами:

- плановые и экстренные госпитализации пациентов;
- отказы в госпитализации без последующей в течение недели повторного направления.

Плановые и экстренные госпитализации описывают группу состояний здоровья пациентов, при котором потребовалась стационарная медицинская помощь без возможности амбулаторного лечения.

Отказы в госпитализации без последующего в течение недели повторного направления описывают группу состояний здоровья пациентов, при котором первоначально лечащим врачом была выявлена потребность оказания стационарной помощи, но в последующем такая необходимость пропала, предполагая амбулаторное лечение.

2.1.2 Описание признаков

При формировании списка признаков были проанализированы как назначаемые лечащими врачами перед направлением на госпитализацию, так и выполняемые в стационаре в первые сутки при поступлении виды лабораторных исследований по вышеуказанным диагнозам. Среди всех просмотренных назначаемых видов исследований были отобраны наиболее часто встречаемые как при экстренной, так и при плановой госпитализации.

Помимо общих показателей здоровья пациента, таких как «вес», «рост», «группа крови» и «резус-принадлежность», и особенных для данных групп диагнозов, таких как «холестерин», «КФК», «ЛДГ», «Тропонин», были отобраны следующие основные назначаемые виды исследований: «общий анализ крови» (ОАК), «биохимический анализ крови» (БАК), «общий анализ мочи» (ОАМ) и «эхокардиография» (ЭхоКГ).

В таблице 1 описаны все виды показателей здоровья, участвующие в качестве признаков объектов модели.

Таблица 1 - Виды показателей здоровья, участвующие при анализе здоровья пациента

Наименование	Тип	Описание/группа вида показателя здоровья
Возраст	Число	Общие данные пациента
Пол	Строка	Общие данные пациента
Группа крови	Строка	Общие данные пациента
Резус-принадлежность	Строка	Общие данные пациента
Рост	Число	Общие данные пациента
Вес	Число	Общие данные пациента
Диагноз	Строка	Предварительный/направительный/основной/заключительный диагноз

Ядерные эритроциты (NRBC)	Число	Общий анализ крови
Ядерные эритроциты (NRBC%)	Число	Общий анализ крови
Эозинофилы (EOS)	Число	Общий анализ крови
Эозинофилы (EOS%)	Число	Общий анализ крови
Гематокрит (HCT)	Число	Общий анализ крови
Базофилы (BASO%)	Число	Общий анализ крови
Базофилы (BASO)	Число	Общий анализ крови
Моноциты (MONO%)	Число	Общий анализ крови
Моноциты (MONO)	Число	Общий анализ крови
Лимфоциты (LYM%)	Число	Общий анализ крови
Лимфоциты (LYM)	Число	Общий анализ крови
Нейтрофилы (NEU%)	Число	Общий анализ крови
Нейтрофилы (NEU)	Число	Общий анализ крови
Тромбоциты (PCT)	Число	Общий анализ крови
Средний объем тромбоцитов (MPV)	Число	Общий анализ крови
Ширина распределений тромбоцитов (PDW)	Число	Общий анализ крови
Коэффициент анизотропии эритроцитов (RDW-CV)	Число	Общий анализ крови
Тромбоциты (PLT)	Число	Общий анализ крови
Средняя концентрация	Число	Общий анализ крови

гемоглобина в эритроцитах (МСНС)		
Среднее содержание гемоглобина в эритроците (МСН)	Число	Общий анализ крови
Средний объем эритроцита (MCV)	Число	Общий анализ крови
Гемоглобин (HGB)	Число	Общий анализ крови
Эритроциты (RBC)	Число	Общий анализ крови
Лейкоциты (WBC)	Число	Общий анализ крови
Глюкоза	Число	Общий анализ крови
Мочевина в крови	Число	Общий анализ крови
Незрелые гранулоциты (IG)	Число	Общий анализ крови
Незрелые гранулоциты (IG%)	Число	Общий анализ крови
МНО	Число	Общий анализ крови
Протромбиновое время	Число	Общий анализ крови
Протромбиновый индекс	Число	Общий анализ крови
Триглицериды	Число	Используется при оценке риска сердечно-сосудистых заболеваний
Фибриноген	Число	Используется при оценке риска сердечно-сосудистых заболеваний
КФК (креатинфосфокиназа)	Число	Используется при оценке риска сердечно-сосудистых заболеваний. Возрастают при остром инфаркте миокарда

МВ КФК (МВ-фракции креатинфосфокиназы)	Число	Используется при оценке риска сердечно-сосудистых заболеваний. Возрастают при остром инфаркте миокарда
ЛДГ (лактатдегидрогеназа)	Число	Используется при оценке риска сердечно-сосудистых заболеваний. Возрастают при остром инфаркте миокарда
CRP	Число	Появляется в организме при воспалительном процессе или уже произошедшем некрозе тканей, поскольку в сыворотке крови здорового человека он содержится в минимальных значениях
Тропонин I	Число	Используется при оценке риска сердечно-сосудистых заболеваний
NT pro BNP Экспресс	Число	Используется в качестве маркера при диагностике сердечной недостаточности. Уровень BNP повышен у пациентов с дисфункцией левого желудочка. При этом содержание BNP в плазме крови достоверно коррелирует с функциональными классами хронической сердечной недостаточности. Определение уровня BNP в плазме крови помогает оценить степень тяжести ХСН и ИБС, прогнозировать

		дальнейшее развитие заболевания, а также оценивать эффект проводимой терапии
ФВ по Teichgolz	Число	Эхокардиография
ФВ по Simpson	Число	Эхокардиография
Трикуспидальная регургитация	Число	Эхокардиография
Трикуспидальный стеноз	Число	Эхокардиография
СДЛА	Число	Эхокардиография
Митральная регургитация	Число	Эхокардиография
КСР	Число	Эхокардиография
КДР	Число	Эхокардиография
Аортальный стеноз	Число	Эхокардиография
Объем ЛП	Число	Эхокардиография
Липопротеиды низкой плотности	Число	Биохимический анализ крови
Липопротеиды высокой плотности	Число	Биохимический анализ крови
Креатинин	Число	Биохимический анализ крови
Калий	Число	Биохимический анализ крови
Скорость клубочковой фильтрации (MDRD)	Число	Биохимический анализ крови
Белок общий	Число	Биохимический анализ крови
Билирубин прямой	Число	Биохимический анализ крови
Билирубин общий	Число	Биохимический анализ крови
АЧТВ	Число	Биохимический анализ крови

АСТ	Число	Биохимический анализ крови
АЛТ	Число	Биохимический анализ крови
Холестерин	Число	Избыток служит риском развития атеросклероза и ишемической болезни сердца. Повышение уровня холестерина (при значительном нарушении функции печени — гипохолестеринемия), триглицеридов, липопротеинов низкой и очень низкой плотности, снижение липопротеинов высокой плотности возможно при ИБС
Глюкоза	Строка	Общий анализ мочи
Кетоны	Строка	Общий анализ мочи
Кислотность	Число	Общий анализ мочи
Лейкоциты	Число	Общий анализ мочи
Цвет	Строка	Общий анализ мочи
Прозрачность	Строка	Общий анализ мочи
МНО	Число	Общий анализ мочи
Удельный вес	Число	Общий анализ мочи
Эпителий плоский	Число	Общий анализ мочи
Эритроциты	Число	Общий анализ мочи

2.1.3 Реализация экспорта набора данных

Перед каждым случаем госпитализации для определения состояния здоровья, лечащим врачом назначается перечень видов лабораторных исследований, который в последствии необходимо пройти пациенту. Отказ от госпитализации в большинстве случаев проводится уже в больнице, куда был направлен пациент.

В МИС «1С.Медицина.Регион» для хранения всех показателей здоровья пациентов существует одноименный регистр сведений «Показатели здоровья» (см. рисунок 1).

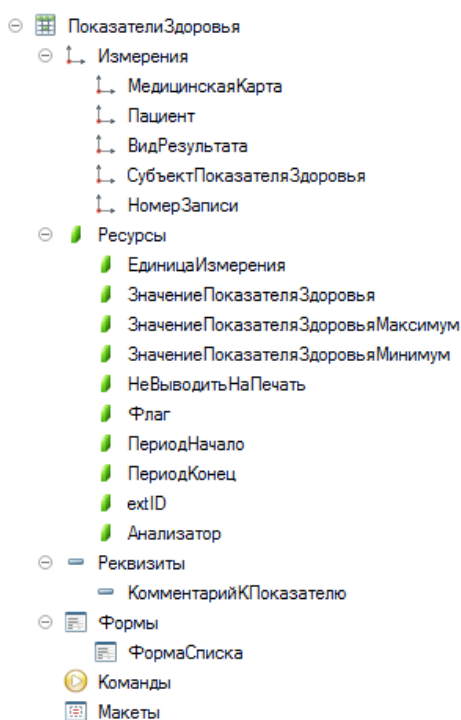


Рисунок 1 - Структура регистра сведений "Показатели здоровья"

Каждое значение вида показателя здоровья обязательно привязывается в трем измерениям: «МедицинскаяКарта», «Пациент» и «ВидРезультата».

Измерение «МедицинскаяКарта» указывает на вид оказываемой медицинской помощи, например, «амбулаторная медицинская карта» или «медицинская карта стационарного больного»; измерение «Пациент» указывает на пациента, для которого выполнялись назначенные лабораторные исследования»; измерение «ВидРезультата» - вид показателя здоровья, которому соответствует значение.

При отборе полного набора значений показателей здоровья необходимо отбирать записи только по измерениям «Пациент» и «ВидРезультата», так как пациент может сдавать показатели здоровья как амбулаторно, так и стационарно.

Для получения результатов проведенных анализов используются несколько лабораторий, что создает проблему реализации механизма

экспорта, так как каждая из них возвращает как свой собственный идентификатор вида показателя здоровья (см. рисунок 2), так и значение. Так, в зависимости от выбранной лаборатории, в числовом значении в качестве разделителя целой и дробной частях может использоваться запятая или точка (см. рисунок 3). Для строкового значения все еще сложнее – может меняться не только окончание, но и язык, регистр, пробелы и т.п.

"ЦЛА027.001",	"БАК_Калий";
"ЦЛА027_001",	"БАК_Калий";
"ЛИС019.001",	"БАК_Калий";
"ЦЛА014.001",	"БАК_Креатинин";
"1006469",	"БАК_Креатинин";
"ЦЛА014_001",	"БАК_Креатинин";
"Креатинин",	"БАК_Креатинин";
"Креатинин крови",	"БАК_Креатинин";
"CREA",	"БАК_Креатинин";
"1006725",	"ЛДГ";
"1006733",	"ЛДГ";
"1006741",	"ЛДГ";
"ЦЛА043_001",	"ЛДГ";
"ЦЛА043.001",	"ЛДГ";
"Лактатдегидрогеназа",	"ЛДГ";
"Креатинфосфокиназа",	"КФК";
"ЦЛА041_001",	"КФК";
"ЦЛА041.001",	"КФК";
"Креатинфосфокиназа МВ",	"МВ_КФК";
"ЦЛА042_001",	"МВ_КФК";
"ЦЛА042.001",	"МВ_КФК";
"ЦЛА020.001",	"БАК_ЛипопротеидыВысокойПлотности";
"ЦЛА020_001",	"БАК_ЛипопротеидыВысокойПлотности";
"ЦЛА021.001",	"БАК_ЛипопротеидыНизкойПлотности";
"ЦЛА021_001",	"БАК_ЛипопротеидыНизкойПлотности";
"Протромбиновое время",	"ПротромбиновоеВремя";
"ЦЛА076_001",	"ПротромбиновоеВремя";
"ЦЛА076.001",	"ПротромбиновоеВремя";
"1000058",	"ПротромбиновоеВремя";
"1000041",	"ПротромбиновоеВремя";
"ЦЛА076.002",	"ПротромбиновыйИндекс";
"ЦЛА076_002",	"ПротромбиновыйИндекс";
"ЦЛА022.001",	"Триглицериды";
"ЦЛА022_001",	"Триглицериды";
"фибриноген",	"фибриноген";
"ЦЛА073_001",	"фибриноген";
"ЦЛА073.001",	"фибриноген";
"ЦЛА019.001",	"Холестерин";
"1009455",	"Холестерин";
"ЦЛА019_001",	"Холестерин";
"Холестерин",	"Холестерин";
"СНОL",	"Холестерин";

Рисунок 2 - Пример наличия различных идентификаторов для соответствия лабораторного исследования

БАК_БилирубинПрямой	БАК_Калий	БАК_Креатинин	БАК_ЛипопротеидыВысокойПлотности	БАК_ЛипопротеидыНизкойПлотности
3,80	NaN	104,1	0,69	3,94
NaN	3,78	96,0	0,97	2,19
3.9	4.57	72	1.16	2.8

Рисунок 3 - Пример наличия разных разделителей в числовых значениях показателей здоровья

В начале реализации инструмента экспорта показателей здоровья были просмотрены и сгруппированы на основании необходимых признаков все имеющиеся в ОКБ 1 виды показателей здоровья из различных лабораторных систем по виду исследования. Для каждой полученной группы было задано «соответствие» - уникальный строковый идентификатор, который явно указывает на выполняемый лабораторный анализ и вид исследования.

Интервал учета показателей здоровья – неделя до даты госпитализации пациента включительно. Данное решение позволит получить наиболее актуальные значения видов показателей здоровья, которые явно указывают на причины случая госпитализации пациента.

Формирование набора данных проводится посредством внешней обработки, при реализации использовались механизмы «Система компоновки данных» (СКД) и «Фоновые задания», которые позволяют как упростить саму разработку, так и снизить нагрузку на СУБД и сервер. Полученные табличные документы выгружаются в формате «.xlsx».

«Система компоновки данных» представляет собой механизм, основанный на декларативном описании отчетов. Он предназначен для построения отчетов, а также вывода информации, имеющей сложную структуру и содержащий произвольный набор таблиц и диаграмм. [5]

«Фоновые задания» – это действия/задачи, производимые программой 1С незаметно для работающего в ней. Отслеживают их исполнение только по журналу регистрации. Особенностью такого задания является то, что при его

выполнении не блокируется работа пользователя, он может продолжать вносить и редактировать данные, даже если процедура запущена. [6]

Текста запросов получения показателей здоровья по случаям госпитализации и отказов в них представлен в приложении 1.

2.1.4 Выгрузка

Выгрузка набора данных проводилась в период с января 2019 года по май 2023 «пачками» - за один раз отбирались записи за определенный интервал, который был равен трем месяцам, в установленное системными администраторами время. Пример формы обработки выгрузки представлен на рисунке 4, а пример файла формата «.xlsx» на рисунке 5.

Всего получилось собрать следующее количество записей:

- 20903 случая госпитализаций пациентов;
- 10258 случаев отказа от госпитализаций пациентов, которые не были повторно госпитализированы в течении следующей недели.

Общее количество признаков – 80.

Данные пациентов были анонимизированы путем выгрузки уникального идентификатора пациента в МИС.

Выгрузка показателей здоровья госпитализированных пациентов (больница) *

Источник выгрузки: Текущая база | Центральная база

Путь к каталогу: Расширение: zip Выгрузить в каталог

Период: 01.02.2023 - 31.05.2023

Сформировать

Идентификатор пациента в РМИС	Дата рождения	Пол	Диагноз	УИД Госпитализации	Канал госпитализации	Дата госпитализации	Дата смерти	CRP	DIA	NT_proBNP_Экстр	SIS	Аортальный
1 069 007	23.03.1969	М	I22.1	0015696a-b351-11ed-893a-005056b6691f	Экстренно	23.02.2023 13:05:20			120		80	
776 726	02.02.1955	М	I20.0	00343eec-cc7b-11ed-8c3a-005056b67ccd	Экстренно	27.03.2023 13:40:01			100		56,6 70	2
713 299	11.01.1949	М	I25.8	00458b6a-b7ff-11ed-982f-005056b6691f	Планово	01.03.2023 12:01:41		1,719	68		561 108	2
541 047	23.03.1936	Ж	I25.1	00572fe4-dcdc-11ed-83fa-005056b67ccd	Планово	17.04.2023 9:54:44		0,232	70		905,9 105	2
13 696	25.12.1957	М	I21.0	0074256e-b026-11ed-9a29-005056b67ccd	Экстренно	19.02.2023 12:20:48	19.02.2023 16:00:00			70		130 2
1 009 107	09.07.1957	Ж	I21.1	00a7e59e-d947-11ed-9042-005056b67ccd	Экстренно	12.04.2023 20:30:37			70		129,9 100	2
300 346	24.09.1960	М	I25.1	00b7c3a0-c223-11ed-8ef3-005056b6691f	Планово	14.03.2023 9:44:50		0,582	90		3 722 130	2
974 357	02.07.1971	М	I25.9	00f4e71c-eaf5-11ed-88d1-005056b67ccd	Планово	05.05.2023 8:28:44				80		130 2
1 006 477	05.04.1953	Ж	I20.0	0146c3d8-b040-11ed-9a29-005056b67ccd	Экстренно	19.02.2023 15:25:55		0,075	84		92,8 130	2
1 150 557	28.09.1968	М	I25.9	0151043c-e04d-11ed-81ea-005056b6691f	Экстренно	21.04.2023 19:01:04		0,17	110		140	
381 606	25.11.1956	М	I25.8	015b56be-a770-11ed-9f5a-005056b6691f	Планово	08.02.2023 10:16:47			70		56,4 140	2
1 021 444	21.07.1954	М	I25.1	015c7f48-de67-11ed-89b8-005056b67ccd	Планово	19.04.2023 9:02:10			70		173 110	2
3 897 064	22.04.1955	М	I25.1	019b51be-c3c8-11ed-89e4-005056b67ccd	Планово	16.03.2023 11:58:28		0,076				
488 172	19.08.1952	М	I25.8	01d711a8-ca31-11ed-9c38-005056b67ccd	Экстренно	24.03.2023 15:45:17	25.03.2023 11:00:00		20,33	51		73
3 875 241	06.01.1962	Ж	I25.9	01db04d0-bf09-11ed-8bd2-005056b6691f	Планово	10.03.2023 11:00:39				60		110
4 010 512	06.01.1963	М	I21.9	021f99c8-d85f-11ed-8eac-005056b6691f	Экстренно	11.04.2023 16:49:48			70		30 110	2
3 753 077	13.07.1961	М	I21.4	024d7ec8-a47b-11ed-90ba-005056b67ccd	Экстренно	04.02.2023 15:59:19						2
515 699	07.01.1964	М	I25.1	0290e3aa-c6d3-11ed-858c-005056b67ccd	Планово	20.03.2023 8:54:48		1,798	62		2 309,5 83	2

Рисунок 4 - Реализованная внешняя обработка выгрузки набора данных

АЧТВ	Вес	КФК	МВ_КФК	Миоглобин	ОАК_Базофилы	ОАК_Гематокри	ОАК_Гемоглоби	ОАК_Лейкоциты	ОАК_Лимфоцит	ОАК_Миелоциты	ОАК_Моноциты
	103.				0,5	41,3	139	6,16	27,3		4,9
	32,8 80 .		18		0,5	44	147	6,46	25,9		11,3
	29,2 152.		13,2			39,8	124	5,05	19,8		5,1
	46,5 87 .				0,8	35,1	126	8,96	37,3		11,9
	37,1				0,8	30,4	90	7,46	19,8		11,7
	95 .		78,1		0,2	41,3	137	13,92	12,8		3,7
	34,4 82 .	112	64,3		0,1	40,9	132	6,89	26,7		5,8
	26,4				0,3	39	132	14,7	7,3		10,1
	73 .		13,1								
	70,9		60,9		0,2	48,8	144	17,37	17,2		6,5
	71 .				0,3	33,7	118	7,92	21,5		9,8
146,4	105.	114	13,0		0,6	43,3	154	5,23	36,5		10,1
	40,2 68 .		16,1		1	32,8	125	9,36	22,9		12,1
	27,9 65 .	47			0,5	46,9	141	11,42	23,6		7,5
	105.		38,8		0,2	41,2	140	9	21,7		3,8
	61 .	56,3									
	65 .	62			0,5	44	142	6,04	31,3		9,1
	36,2 65 .		11,4		0,4	37,3	125	9,58	21,5		5,1
	104.		36,8		0,5	46,3	146	8,64	33,3		9,6
	35 89 .	175,7			0,4	33,6	110	8,32	14,5		5,3
			26,7		0,3	42,4	128	10,18	13,6		1,4
	116 90 .		106,2		0,6	38,9	135	13,95	9,9		3,4
	30,2 103.	146,8	26,1		0,7	36,5	138	7,5	26,7		3,5
			50,5		0,2	45,1	153	10,04	25,3		5,8

Рисунок 5 - Пример заполнения данными файла ".xlsx"

2.2 Сравнительный анализ алгоритмов машинного обучения

Исходя из вышеописанных целей, задачу прогнозирования можно свести к бинарной классификации со следующими описаниями значений целевой переменной:

- 1 – существует риск ухудшения состояния здоровья;
- 0 – риск ухудшения состояния здоровья незначительный или отсутствует.

Бинарная классификация используется для решения вопроса о принадлежности некоторого объекта к одному из двух классов. Как показывает практика, множество задач классификации в Data Mining может быть сведено к бинарным. При их использовании удается упростить модель и снять некоторые ограничения, связанные с большим числом возможных состояний выходного значения. [7]

Наиболее часто в качестве используемых на практике моделей бинарной классификации или прогнозирования можно встретить следующие:

- логистическая регрессия;
- К-ближайшие соседи;
- случайный лес;
- одноклассовая классификация SVM;
- градиентный бустинг;

2.2.1 Логистическая регрессия

Логистическая регрессия (Logistic Regression) — это модель линейной классификации, которая используется для моделирования двоичной целевой переменной. Он используется для прогнозирования вероятности того, что событие произойдет. Для определения класса на основании полученного значения задается некий порог P так, что если $p \geq P$, то на выходе будет 1, а иначе 0. Используемая регрессией сигмоидальная функция (см. рисунок 5) ограничена двумя горизонтальными асимптотами, из-за чего значение вероятности находится в промежутке $[0, 1]$. [8]

$$f(x) = \frac{1}{1 + e^{-x}}$$

Рисунок 6 - Сигмоидальная функция

К преимуществам логистической регрессии можно отнести:

- предоставляет оценку вероятности принадлежности объекта к классу 1;
- проста в изучении и использовании.

Недостатки:

- подходит только для задач, где данные линейно зависимы;

2.2.2 K-ближайшие соседи.

K-ближайшие соседи (k-nearest neighbors, KNN) — популярный алгоритм классификации, который используется в разных типах задач машинного обучения. Является нелинейным классификатором, который предсказывает, к какому классу принадлежит новая точка тестовых данных, путем определения класса k ближайших соседей. Выбирается k ближайших соседей на основе евклидова расстояния. Среди этих k соседей подсчитывается количество точек данных в каждой категории, и новая точка данных назначается той категории, в которой у нас больше всего соседей. [9]

Интерпретировать результат работы метода можно следующим образом: объект относится к тому же классу, что и большинство его ближайших соседей.

К преимуществам данной модели можно отнести:

- простота в понимании и реализации;
- устойчивость к выбросам и аномальным значениям;
- результаты работы алгоритма легко поддаются интерпретации.

Недостатки:

- является достаточно затратным в вычислительном плане;
- необходимо вручную выбрать количество соседей k.

2.2.3 Случайный лес

Случайный лес (Random forest, RF) представляет из себя ансамбль независимых деревьев решений. Каждое дерево обучается независимо от других на случайной выборке данных. Это помогает сделать модель более надежной, чем одна модель дерева решений, и меньшая вероятность получить переобучение. [10]

Обычно, в RF есть два параметра – количество деревьев-оценщиков и количество элементов в случайной листе.

К достоинствам данной модели можно отнести:

- способность эффективно обрабатывать данные с большим числом признаков и классов;
- случайные леса очень гибки и обладают очень высокой точностью.

Недостатки:

- большой размер получающихся моделей;
- построение леса сложнее и отнимает больше времени.

2.2.4 Одноклассовая классификация SVM

SVM (support vector machine, метод опорных векторов) - один из наиболее популярных методов обучения, который применяется для решения задач классификации и регрессии. Основная идея метода заключается в построении гиперплоскости, разделяющей объекты выборки оптимальным способом. Алгоритм работает в предположении, что чем больше расстояние (зазор) между разделяющей гиперплоскостью и объектами разделяемых классов, тем меньше будет средняя ошибка классификатора. [11]

Одноклассовая классификация SVM (OneClassSVM) используется при решении задач выявления «новизны»/«аномалий» или «выбросов». Как и выброс «новый объект» — это объект, который отличается по своим свойствам от объектов (обучающей) выборки. Но в отличие от выброса, его в самой выборке пока нет.

Новизна, как правило, появляется в результате принципиально нового поведения объекта. Если наши объекты – описания работы системы, то после проникновения в неё вируса объекты становятся «новизной». [12]

Пример задач с использованием поиска «новизны»:

- обнаружение подозрительных банковских операций;
- обнаружение неполадок в механизмах по показаниям датчиков;
- медицинская диагностика.

Если подразумевать, что «стандартные» данные – это пациенты с наличием риска, а «новизна» - пациенты с его отсутствием, то OneClassSVM становится хорошим выбором среди прочих алгоритмов машинного обучения, так как при обучении модель «затачивается» под обучающую выборку, которую составляют объекты одного единственного класса.

2.2.5 Градиентный бустинг

Градиентный бустинг – алгоритм машинного обучения для решения задач классификации и регрессии. Предсказание строится на основе ансамбля различных моделей, которые в основном являются деревьями решений. Каждая последующая модель ансамбля сводит к минимуму ошибку предыдущей. Считается одной из самых эффективных инструментов для решения задач машинного обучения. [13]

Большое количество реальных задач классификации решаются посредством ансамбля деревьев, и среди них градиентный бустинг показывает себя часто лучше своих аналогов.

Преимущества:

- алгоритм работает с любыми функциями потерь;
- предсказания в среднем лучше, чем у других алгоритмов;
- самостоятельно может справляться с пропущенными данными.

Недостатки:

- алгоритм крайне чувствителен к выбросам;
- вычисления могут занять много времени;

- модель будет склонна к переобучению при слишком большом количестве деревьев.

2.2.6 Вывод

При сравнительном анализе возможностей, применений и отзывов рассмотренных моделей, было решено остановиться на алгоритме «градиентного бустинга».

2.3 Реализация предварительной обработки данных

Помимо признаков, в наборе данных присутствуют «служебные» поля, которые в дальнейшем либо понадобятся при разработке интеграции с МИС «1С:Медицина.Регион» («уникальный идентификатор РМИС», «УИД госпитализации»), либо на основании которых будут формироваться другие признаки, например, расчет возраста на основании «даты госпитализации»/«даты отказа» и «даты рождения».

По причине наличия нескольких лабораторий, получаемые МИС данные по одному и тому же виду лабораторного исследования могут сильно различаться регистром, языком написания, наличием различных знаков пунктуации, дополнительных цифр, пробелов и прочим.

В зависимости от типа данных, «количественное» или «категориальное» значение, необходимо подобрать наиболее оптимальный способ обработки значений для каждого признака из набора данных.

2.3.1 Используемые технические средства и программы

При реализации использовались следующие технические средства и программы:

- язык программирования Python 3.9.7;
- среда разработки Visual Studio Code.

Используемые библиотеки для Python:

- Pandas (v 2.0.1);
- Numpy (v 1.24.2);
- Matplotlib (v 3.7);

- Glob;
- Textdistance (v 4.5.0).

Набор данных изначально представлен двумя группами файлов для дальнейшего установления меток каждой в отдельности:

- планоно или экстренно госпитализированные пациенты;
- пациенты с отказом в госпитализации без повторного направления в течение последующей недели;

Каждая группа загружается отдельно – сначала «госпитализированные случаи», потом «отказы». Для загрузки файлов формата «.xlsx» использовались библиотеки «Glob» и «Pandas», каждый набор представлен в виде объекта «DataFrame» библиотеки «Pandas».

2.3.2 Процесс предобработки данных

В полученных наборах удаляются пустые строки и колонки, в соответствующих полях проводится приведение строковых значений к датам и последующий расчет признака «Возраст».

Пример полученного набора данных госпитализированных пациентов представлен на рисунке 6.

БАК_ЛипопротендыВысокойПлотности	БАК_ЛипопротендыНизкойПлотности	БАК_СкоростьКлубочковойФильтрацииMDRD	БАК_ТропонинI	Вес	ГруппаКрови	КФК	ЛДГ	МВ_КФК	МНО	ОАК_Базофилы_BASO
0.69	3.94	67.19	0.003	111.	B(III) третья	NaN	NaN	15.5	1.00	0,040
0.97	2.19	76.87	0.001	105.	O(I) первая	60.0	NaN	11.8	1.10	0,050
1.16	2.8	106.7	0.004	93.	A?B(V) четвертая со слабым антигеном "A"	66	NaN	16.8	1.05	0,03
NaN	NaN	NaN	NaN	62.	NaN	89.0	425.0	11.2	1.12	0,04
NaN	NaN	NaN	NaN	69.	A(II) вторая	NaN	NaN	NaN	1.08	NaN
0.67	2.34	25.35	0.014	126.	A(II) вторая	150.0	NaN	20.3	2.24	0,010
0.94	1.98	63.90	0.003	61	O(I) первая	NaN	NaN	12.6	4.37	0,030
1.01	1.99	85.43	NaN	82.	O(I) первая	68	NaN	22.4	NaN	NaN
NaN	3.26	NaN	0.001	65.	B(III) третья	NaN	141	NaN	1.03	0,01
NaN	NaN	NaN	NaN	76.	NaN	121.0	431.0	NaN	1.27	0,03

Рисунок 7 - Пример записей набора данных госпитализированных пациентов

На изображении видно как наличие пропусков, так и различия в написании значений. Среди всего набора данных, сильно выделяются поля

«Группа крови» и «Резус-принадлежность». В данных колонках есть проблема не только с регистром, пробелами или спец. символами, но и языке – некоторые значения написаны на русском, другие – на английском, а иногда – на обоих.

#	Column	Non-Null Count	Dtype
0	Идентификатор пациента в РМИС	9924 non-null	float64
1	Дата рождения	10258 non-null	object
2	Пол	10258 non-null	object
3	Диагноз	10258 non-null	object
4	УИД отказа	10258 non-null	object
5	Дата отказа	10258 non-null	object
6	CRP	1907 non-null	object
7	DIA	3726 non-null	float64
8	NT_proBNP_Экспресс	894 non-null	object
9	SIS	3726 non-null	float64
10	БАК_АЛТ	4457 non-null	object
11	БАК_АСТ	4454 non-null	object
12	БАК_АЧТВ	3997 non-null	object
13	БАК_БелокОбщий	3645 non-null	object
14	БАК_БилирубинОбщий	4159 non-null	object
15	БАК_БилирубинПрямой	3819 non-null	object
16	БАК_Калий	5497 non-null	object
17	БАК_Креатинин	10046 non-null	object
18	БАК_ЛипопротеидыВысокойПлотности	2409 non-null	object
19	БАК_ЛипопротеидыНизкойПлотности	2478 non-null	object
20	БАК_СкоростьКлубочковойФильтрацииMDRD	6960 non-null	object
21	БАК_ТропонинI	9563 non-null	object
22	Вес	3899 non-null	object
23	ГруппаКрови	2955 non-null	object
24	КФК	1735 non-null	object
25	ЛДГ	1186 non-null	object
26	МВ_КФК	9382 non-null	object
27	МНО	4254 non-null	object
28	ОАК_Базофилы_BASO	9909 non-null	object
29	ОАК_Базофилы_BASO%	9906 non-null	object
30	ОАК_Гематокрит_HCT	10169 non-null	object
31	ОАК_Гемоглобин_HGB	10168 non-null	object
32	ОАК_Глюкоза	9317 non-null	object
33	ОАК_КоэффАнизотрЭритр_RDWCV	10159 non-null	object
34	ОАК_Лейкоциты_WBC	10168 non-null	object
35	ОАК_Лимфоциты_LYM	10162 non-null	object
36	ОАК_Лимфоциты_LYM%	10163 non-null	object
37	ОАК_Моноциты_MONO	10161 non-null	object
38	ОАК_Моноциты_MONO%	10161 non-null	object
39	ОАК_Мочевина	3839 non-null	object
40	ОАК_НезрелыеГранулоциты_IG	8766 non-null	object
41	ОАК_НезрелыеГранулоциты_IG%	8797 non-null	object
42	ОАК_Нейтрофилы_NEU	10160 non-null	object
43	ОАК_Нейтрофилы_NEU%	10161 non-null	object

Рисунок 8 - Описание полей признаков объекта DataFrame с отказами в госпитализации

На рисунке 7 можно наблюдать большое количество пропусков в различных полях объекта DataFrame. В данном случае, нельзя с

уверенностью сказать, что пропуски вызваны ошибками пользователя или возникли при выгрузке, так как часто лечащие врачи амбулатории направляют пациентов не на все возможные виды лабораторных исследований, а только на определенные, особенно в том случае, когда изначально известен диагноз и история лечения.

Для последующей предобработки данных из полученных объектов удаляются служебные поля, такие как:

- Для «случаев госпитализации»:
 - «Идентификатор пациента в РМИС»;
 - «Дата госпитализации»;
 - «Дата рождения»;
 - «Канал госпитализации»;
 - «УИД госпитализации».
- Для «отказов в госпитализации»:
 - «Идентификатор пациента в РМИС»;
 - «Дата отказа»;
 - «Дата рождения»;
 - «УИД отказа».

В наборах добавляется поле «mark», для «случаев госпитализаций» устанавливается значение «1», а наборе «отказов в госпитализации» - «0».

После удаления «служебных» полей и установления меток проводится объединение наборов в один объект «DataFrame». Полный список полученных полей и количества заполненных в них значений представлен на рисунке 8.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 31161 entries, 0 to 31160
Data columns (total 82 columns):
#   Column                                                                 Non-Null Count  Dtype
---  -
0   Пол                                                                    31161 non-null  object
1   Диагноз                                                                31161 non-null  object
2   CRP                                                                    8417 non-null   object
3   DIA                                                                    17636 non-null  float64
4   NT_проBNP_Экспресс                                                    5358 non-null   object
5   SIS                                                                    17634 non-null  float64
6   БАК_АЛТ                                                                21462 non-null  object
7   БАК_АСТ                                                                21457 non-null  object
8   БАК_АЧТВ                                                                19657 non-null  object
9   БАК_БелокОбщий                                                       18627 non-null  object
10  БАК_БилирубинОбщий                                                    18983 non-null  object
11  БАК_БилирубинПрямой                                                  18549 non-null  object
12  БАК_Калий                                                             18598 non-null  object
13  БАК_Креатинин                                                         28675 non-null  object
14  БАК_ЛипопротеидыВысокойПлотности                                    16367 non-null  object
15  БАК_ЛипопротеидыНизкойПлотности                                    16727 non-null  object
16  БАК_СкоростьКлубочковойФильтрацииMDRD                              20125 non-null  object
17  БАК_ТропонинI                                                         24537 non-null  object
18  Вес                                                                    21023 non-null  object
19  ГруппаКрови                                                           21196 non-null  object
20  КФК                                                                    7739 non-null   object
21  ЛДГ                                                                    4884 non-null   object
22  МВ_КФК                                                                25202 non-null  object
23  МНО                                                                    21352 non-null  object
24  ОАК_Базофилы_BASO                                                    27213 non-null  object
25  ОАК_Базофилы_BASO%                                                  27164 non-null  object
26  ОАК_Гематокрит_HCT                                                  28407 non-null  object
27  ОАК_Гемоглобин_HGB                                                  28275 non-null  object
28  ОАК_Глюкоза                                                           27497 non-null  object
29  ОАК_КоэффициентЭритроцитР_RDWCV                                    28247 non-null  object
30  ОАК_Лейкоциты_WBC                                                    28275 non-null  object
31  ОАК_Лимфоциты_LYM                                                    28248 non-null  object
32  ОАК_Лимфоциты_LYM%                                                  28251 non-null  object
33  ОАК_Моноциты_MONO                                                    28242 non-null  object
34  ОАК_Моноциты_MONO%                                                  28244 non-null  object
35  ОАК_Мочевина                                                         20279 non-null  object
36  ОАК_НезрелыеГранулоциты_IG                                          24426 non-null  object
37  ОАК_НезрелыеГранулоциты_IG%                                         24662 non-null  object
38  ОАК_Нейтрофилы_NEU                                                  28241 non-null  object
39  ОАК_Нейтрофилы_NEU%                                                 28250 non-null  object
40  ОАК_Сред_конц_гемоглобина_в_эр_МСНС                                28258 non-null  object
41  ОАК_Сред_сод_гемоглобина_эритроците_МСН                            28259 non-null  object
42  ОАК_Средний_объем_эритроцита_MCV                                    28259 non-null  object
43  ОАК_СреднийОбъемТромбоцитов_MPV                                    28114 non-null  object
44  ОАК_Тромбоциты_PCT                                                  28106 non-null  object
45  ОАК_Тромбоциты_PLT                                                  28273 non-null  object
46  ОАК_ШиринаРаспредТромбоцитов_PDW                                    28114 non-null  object
47  ОАК_Эозинофилы_EOS                                                  175 non-null    object
48  ОАК_Эозинофилы_EOS%                                                 26289 non-null  object
49  ОАК_Эритроциты_RBC                                                  28275 non-null  object
50  ОАМ_Глюкоза                                                           18165 non-null  object
51  ОАМ_Кетоны                                                            18838 non-null  object
52  ОАМ_Кислотность                                                       19365 non-null  object
53  ОАМ_Лейкоциты                                                         19277 non-null  object
54  ОАМ_Прозрачность                                                      19364 non-null  object
55  ОАМ_УдельныйВес                                                       19353 non-null  object
56  ОАМ_Цвет                                                               19366 non-null  object
57  ОАМ_ЭпителийПлоский                                                  19261 non-null  object
58  ОАМ_Эритроциты                                                         17664 non-null  object
59  ПротромбиновоеВремя                                                 21150 non-null  object
60  ПротромбиновыйИндекс                                                 21161 non-null  object
61  РезусПринадлежность                                                  21054 non-null  object
62  РезусФенотипирование                                                 15963 non-null  object
63  Рост                                                                    20985 non-null  object
64  Триглицериды                                                         16386 non-null  object

```

65	Фибриноген	21162	non-null	object
66	Холестерин	21688	non-null	object
67	ОАК_Ядерные_эритроциты_NRBC	831	non-null	float64
68	ОАК_Ядерные_эритроциты_NRBC%	818	non-null	float64
69	АортальныйСтеноз	6999	non-null	float64
70	КДР	6973	non-null	float64
71	КСР	6966	non-null	float64
72	МитральнаяРегургитация	6400	non-null	object
73	ОбъемЛП	4732	non-null	float64
74	СДЛА	6185	non-null	float64
75	ТрикуспидальнаяРегургитация	6136	non-null	object
76	ТрикуспидальныйСтеноз	6999	non-null	float64
77	ФBSimpson	5284	non-null	float64
78	ФVTeichgolz	6687	non-null	float64
79	Возраст	31161	non-null	float64

Рисунок 9 - Список признаков набора данных и их описание

Весь дальнейший процесс предобработки набора данных был разделен на следующие этапы:

1. приведение строковых значений к числовым в соответствующих полях
2. заполнение пропусков;
3. приведение значений категориальных значений к формам без окончаний (стемминг);
4. преобразование категориальных значений к числовым (бинарное кодирование категориальных значений);

Для приведения строковых значений к числовым в соответствующих полях была написана функция, принимающая в качестве параметров объект «DataFrame» и список колонок и применяющая для каждого значения колонок регулярные выражения (см. рисунок 9).

2	CRP	8417	non-null	float64
3	DIA	17636	non-null	float64
4	NT_proBNP_Экспресс	5358	non-null	float64
5	SIS	17634	non-null	float64
6	БАК_АЛТ	21462	non-null	float64
7	БАК_АСТ	21457	non-null	float64
8	БАК_АЧТВ	19654	non-null	float64
9	БАК_БелокОбщий	18627	non-null	float64
10	БАК_БилирубиОбщий	18983	non-null	float64
11	БАК_БилирубинПрямой	18549	non-null	float64
12	БАК_Калий	18598	non-null	float64
13	БАК_Креатинин	28675	non-null	float64
14	БАК_ЛипопротеидыВысокойПлотности	16367	non-null	float64
15	БАК_ЛипопротеидыНизкойПлотности	16727	non-null	float64
16	БАК_СкоростьКлубочковойФильтрацииMDRD	20125	non-null	float64
17	БАК_ТропонинI	24537	non-null	float64
18	Вес	21023	non-null	float64
19	КФК	7739	non-null	float64
20	ЛДГ	4884	non-null	float64
21	МВ_КФК	25202	non-null	float64
22	МНО	21352	non-null	float64
23	ОАК_Базофилы_BASO	27213	non-null	float64
24	ОАК_Базофилы_BASO%	27164	non-null	float64
25	ОАК_Гематокрит_HCT	28407	non-null	float64
26	ОАК_Гемоглобин_HGB	28275	non-null	float64
27	ОАК_Глюкоза	27497	non-null	float64
28	ОАК_КоэффАнизотрЭритр_RDWCV	28247	non-null	float64
29	ОАК_Лейкоциты_WBC	28275	non-null	float64
30	ОАК_Лимфоциты_LYM	28248	non-null	float64
31	ОАК_Лимфоциты_LYM%	28251	non-null	float64
32	ОАК_Моноциты_MONO	28242	non-null	float64
33	ОАК_Моноциты_MONO%	28242	non-null	float64
34	ОАК_Мочевина	20279	non-null	float64
35	ОАК_НезрелыеГранулоциты_IG	24426	non-null	float64
36	ОАК_НезрелыеГранулоциты_IG%	24662	non-null	float64

Рисунок 10 - Преобразованные к числам поля DataFrame

Заполнение пропусков, в виду выбранного алгоритма машинного обучения, было решено реализовать через установление «выброса», равным «-999999». Таким образом, деревья смогут явно определять случаи с неполным списком показателей здоровья и соответствующе обрабатывать (см. рисунок 10).

БАК_АЛТ	БАК_АСТ	БАК_АЧТВ	БАК_БелокОбщий	БАК_БилирубинОбщий	БАК_БилирубинПрямой	БАК_Калий	БАК_Креатинин
29.0	16.4	29.5	69.8	11.5	3.8	-999999.00	104.1
69.3	23.9	75.3	71.2	7.7	-999999.0	3.78	96.0
32.0	18.3	29.0	68.0	10.4	3.9	4.57	72.0
20.2	25.1	33.4	73.0	37.1	13.9	4.31	156.0
-999999.0	-999999.0	30.7	-999999.0	-999999.0	-999999.0	-999999.00	-999999.0
99.5	104.2	33.1	58.3	30.5	17.7	5.20	246.0
16.3	13.4	66.8	59.7	30.3	10.7	5.19	108.1
-999999.0	-999999.0	-999999.0	-999999.0	15.8	4.9	5.01	84.3
17.6	14.5	31.3	70.8	3.6	1.5	3.91	62.0
34.2	23.8	34.4	69.7	18.1	5.2	5.00	88.0

Рисунок 11 - Заполнение пропусков выбросом

Стемминг категориальных значений проводится для двух полей – «ОАМ_Прозрачность» и «ОАМ_Цвет», которые в виду различных

лабораторий имеют разные окончания и написание. Была написана функция, которая использует «расстояние Левенштайна» для поиска заранее заданных оснований слов и преобразует значения к найденному (см. рисунок 11).

ОАМ_Прозрачность	ОАМ_УдельныйВес	ОАМ_Цвет
прозрач	1.012	желт
прозрач	1.034	соломенно-желт
прозрач	1030.000	соломенно-желт
прозрач	1.016	соломенно-желт
NaN	NaN	NaN
прозрач	1.026	желт
прозрач	1.020	желт
NaN	NaN	NaN
прозрач	1020.000	NaN
прозрач	1.000	NaN

Рисунок 12 - Обработка значений полей "ОАМ_Прозрачность" и "ОАМ_Цвет"

С помощью объекта «OneHotEncoding» проводится бинарное кодирование категориальных признаков, после чего выполняется объединение полученного набора с числовыми признаками с помощью функции «concat» библиотеки «Pandas».

2.4 Реализация модели градиентного бустинга

2.4.1 Используемые технические средства и программы

При реализации модели прогнозирования использовались программные средства разработки:

- язык программирования Python 3.9.7;
- среда разработки Visual Studio Code;
- библиотека Sklearn, Matplotlib.

Полученный ранее набор данных был разделен на «набор признаков» и «целевая переменная». Значения «целевой переменной» принимают два

значения: 1 – есть риск ухудшения состояния, 0 – нет риска ухудшения состояния.

Наборы данных были разделены на «обучающую» и «тестовую» выборки в соотношении 80/20 – 24928 и 6233 записей соответственно.

2.4.2 Подбор гиперпараметров модели

Для подбора гиперпараметров был использован объект «GridSearch» библиотеки «Sklearn» (см. рисунок 12). Заранее был определен список значений гиперпараметров модели, среди которых будет проводиться поиск такой комбинации, при которой модель будет показывать лучшую оценку.

```
params_grid = {  
    "n_estimators": [100, 150, 200],  
    "max_depth": [2, 3, 4, 5, 6, 7],  
    "subsample": [0.6, 0.9, 1],  
    "learning_rate": [0.6, 0.8, 0.9, 1],  
    'min_samples_leaf': [8, 10, 12],  
    'min_samples_split': [2, 6, 10, 12]  
}
```

Рисунок 13 - Рассматриваемые значения гиперпараметров модели

На основании результатов работы «GridSearch» был определен набор гиперпараметров, при котором точность модели достигает максимума (см. рисунок 13).

```
{'subsample': 1, 'n_estimators': 100, 'min_samples_split': 6, 'min_samples_leaf': 10, 'max_depth': 7, 'learning_rate': 0.9}
```

Рисунок 14 - Итоговый набор гиперпараметров модели

2.4.3 Оценки качества обученной модели прогнозирования

После обучения и получения тестовых прогнозов, для оценки качества модели использовалась матрица ошибок классификации и AUC-ROC. Выбор данных метрик в основном обусловлен «дизбалансом» классов – количество объектов в классе «1» в 4-5 раз больше, чем в другом классе.

Матрица ошибок используется как метрика качества классификации 2 и более классов с целью сопоставления прогнозируемых и фактических значений [14].

В случае с бинарной классификацией, в матрице ошибок будет 4 ячейки:

- «True Positive» - положительный прогноз совпал с положительным фактическим значением;
- «False Positive» - отрицательный прогноз не совпал с положительным фактическим значением;
- «False Negative» - положительный прогноз не совпал с отрицательным фактическим значением;
- «True Negative» - отрицательный прогноз совпал с отрицательным фактическим значением.

Введем определение двум прогнозируемым классам:

- класс «1» – положительный. Обозначают пациентов с риском ухудшения состояния здоровья;
- класс «0» - отрицательный. Обозначают пациентов без риска ухудшения состояния здоровья.

		Предсказанные значения	
		Положительный (класс «1»)	Отрицательный (класс «0»)
Реальные значения	положительный	TP (истинно-положительные значения)	FP (ложноположительные)
	отрицательный	FN (ложно-отрицательный)	TN (истинно-отрицательный)

Результат тестирования модели в виде матрицы ошибок представлены на рисунках 15.

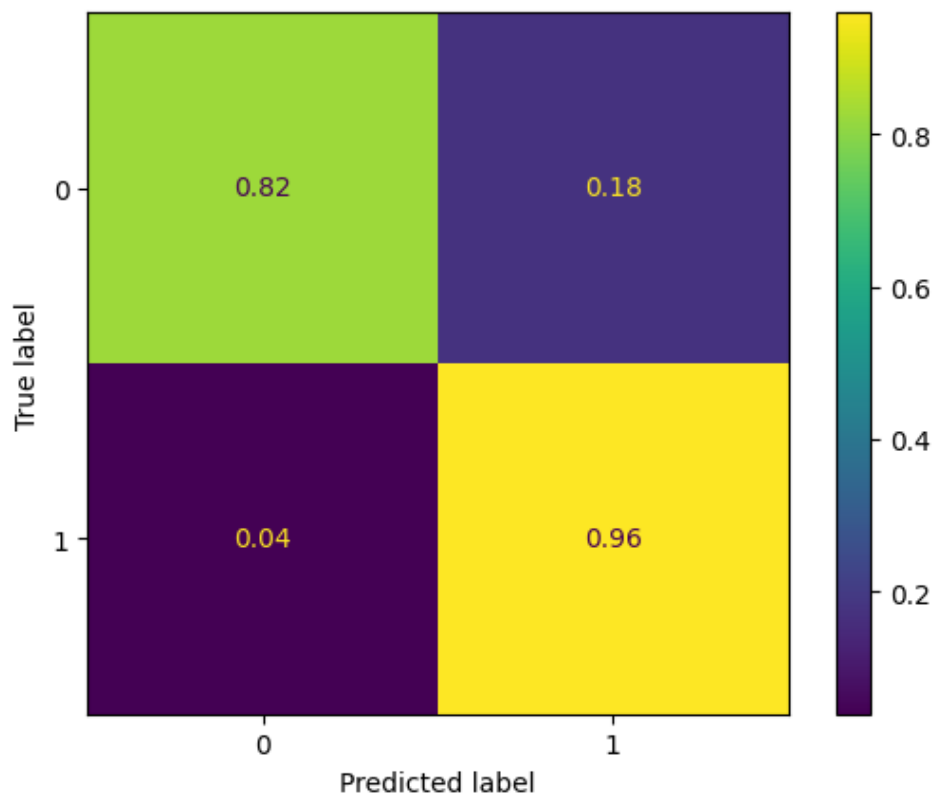


Рисунок 15 - Матрица ошибок

На рисунке можно отметить отличную способность классифицировать пациентов с риском ухудшения состояния здоровья – 0.96. Что касается класса пациентов без риска – 0.82 не такое плохое значение, так как прогнозирование риска все равно в последствии будет обработано лечащим врачом.

Дополнительно, качество работы модели было проверено с помощью метрики AUC-ROC. AUC-ROC (площадь под ROC-кривой) – метрика, которая описывает способность классификатора различать классы. Чем выше площадь под кривой, тем лучше производительность модели.

Результат проверки качества работы классификатора с помощью метрики AUC-ROC представлена на рисунке 16.

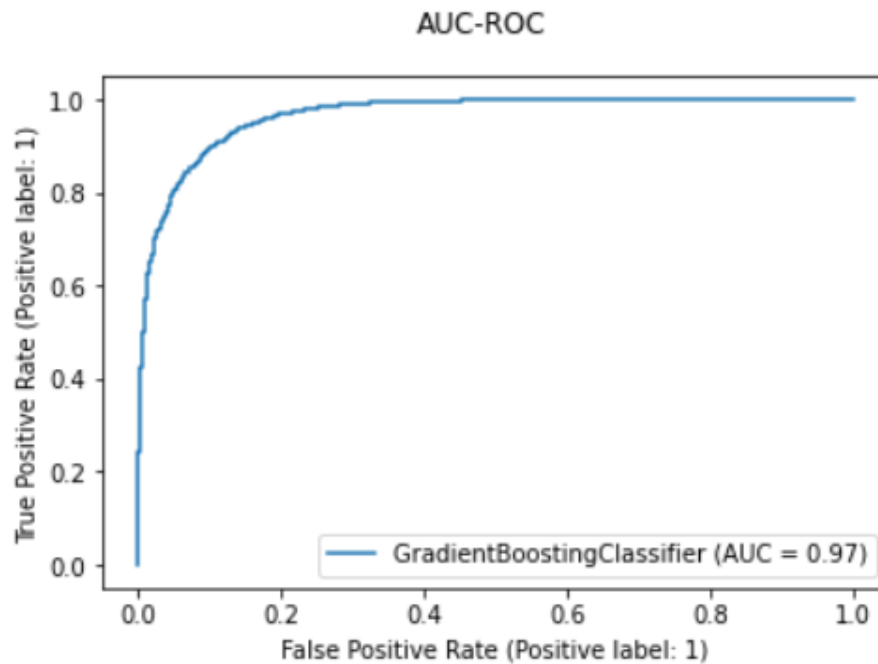


Рисунок 16 - AUC-ROC

Значение площади под кривой ROC = 0.97 показывает отличную возможность модели различать классы «с риском» и «без риска» ухудшения состояния здоровья пациентов.

2.5 Вывод

Таким образом, на основании случаев госпитализации и отказов от госпитализаций с 2019 по май 2023 года по диагнозам ИБС, ХСН и ОКС был сформирован набор данных, признаками которого являются показатели здоровья, которые учувствуют в принятии решения лечащим врачом о дальнейших действиях в плане оказания медицинской помощи.

Для выгруженного набора данных был описан и реализован процесс предобработки, включающий чистку пустых строк и колонок, заполнение пропущенных значений, выделение основы слова у категориальных данных и дальнейшее бинарное кодирование полученных значений.

Реализована модель градиентного бустинга, выполняющая классификацию по двум классам – «есть риск» и «нет риска» ухудшения состояния здоровья пациентов. Подобраны гиперпараметры, оценена работа классификатора.

3. РЕАЛИЗАЦИЯ ВЕБ-СЕРВИСА

3.1 Используемые технические средства и программы

При реализации модели прогнозирования использовались программные средства разработки:

- язык программирования Python 3.9.7;
- среда разработки Visual Studio Code;
- библиотеки Flask, Pickle.

3.2 Разработка

Разработка серверной части производилась на web-фреймворке Flask с использованием FLASK-RESTful'a на языке программирования Python.

Документирование всех запросов будет происходить в Swagger - интерфейс, в котором удобно отображается спецификация по запросам для разработчика клиентской части.

Общение клиентской и серверной части происходит посредством HTTP/HTTPS запросов. Входные и выходные данные будут поставляться в формате JSON.

Доступ к обученной модели и предобработке данных осуществляется через библиотеку Pickle, которая используется для сериализации и десериализации объектов для повторного использования и версионирования.

Проверку HTTP запросов осуществлялась с помощью интерфейса Swagger и с помощью Postman.

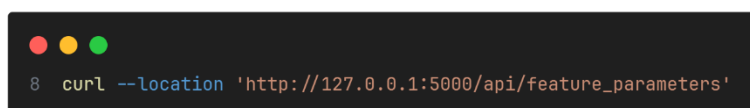
Postman – инструмент для тестирования API приложений. Из преимуществ: сохраняет в памяти устройства коллекции запросов, есть возможность выгрузить запрос на другие языки программирования и специальные http-клиенты (например, curl).

Выбор Flask'a обусловлен тем, что для MVP-версии было реализовано несколько методов без связи с базой данных. Список методов представлен в таблице 2.

Таблица 2 - Спецификация веб-сервиса

API endpoint	Тип запроса	Описание метода	Параметры запроса
predictionOfHealthDeterioration /api/SSZ/feature_parameters	GET	Возвращает список признаков - соответствий видов показателей здоровья разбитые	-
predictionOfHealthDeterioration /api/SZZ/predict	POST	Предсказывает класс и вероятность уверенности в данном класс для набора данных	Виды показателей здоровья, соответствующие результату запроса «/api/feature_parameters»

Тестирование метода получения списка параметров прогнозирования «predictionOfHealthDeterioration/api/SSZ/feature_parameters» представлено на рисунках 17 и 18.



```
8 curl --location 'http://127.0.0.1:5000/api/feature_parameters'
```

Рисунок 17 - Запрос получения списка признаков

```

8  curl --location 'http://127.0.0.1:5000/api/predict' \
9  --header 'Content-Type: application/json' \
10 --data '{
11     "Пол": "М",
12     "Диагноз": "I50.0",
13     "CRP": null,
14     "DIA": 40.0,
15     "NT_проBNP_Экспресс": "22 037,00",
16     "SIS": 80.0,
17     "БАК_АЛТ": "99,5",
18     "БАК_АСТ": "104,2",
19     "БАК_АЧТВ": "33,1",
20     "БАК_БелокОбщий": "58,3",
21     "БАК_БилирубинОбщий": "30,5",
22     "БАК_БилирубинПрямой": "17,70",
23     "БАК_Калий": "5,20",
24     "БАК_Креатинин": "246,0",
25     "БАК_ЛипопротеидыВысокойПлотности": "0,67",
26     "БАК_ЛипопротеидыНизкойПлотности": "2,34",
27     "БАК_СкоростьКлубочковойФильтрацииMDRD": "25,35",
28     "БАК_ТропонинI": "0,014",
29     "Вес": "126.",
30     "КФК": "150,0",
31     "ЛДГ": null,
32     "МВ_КФК": "20,3",
33     "МНО": "2,24",
34     "ОАК_Базофилы_BAS0": "0,010",
35     "ОАК_Базофилы_BAS0%": "0,1",
36     "ОАК_Гематокрит_НСТ": "50,9",
37     "ОАК_Гемоглобин_HGB": 154.0,
38     "ОАК_Глюкоза": "5,88",
39     "ОАК_КoeffАнизотрЭритр_RDWCV": "14,1",
40     "ОАК_Лейкоциты_ВBC": "11,29",
41     "ОАК_Лимфоциты_LYM": "2,27",
42     "ОАК_Лимфоциты_LYM%": "20,1",
43     "ОАК_Моноциты_MONO": "0,44",
44     "ОАК_Моноциты_MONO%": "3,9",
45     "ОАК_Мочевина": "20,00",
46     "ОАК_НезрелыеГранулоциты_IG": "0,02",
47     "ОАК_НезрелыеГранулоциты_IG%": "0,20",
48     "ОАК_Нейтрофилы_NEU": "8,57",
49     "ОАК_Нейтрофилы_NEU%": "75,9",
50     "ОАК_Сред_конц_гемоглобина_в_эр_MCHC": 337.0,
51     "ОАК_Сред_сод_гемоглобина_эритроците_MCH": "30,4",
52     "ОАК_Средний_объем_эритроцита_MCV": "90,1",
53     "ОАК_СреднийОбъемТромбоцитов_MPV": "12,9",
54     "ОАК_Тромбокриты_PCT": "0,2",
55     "ОАК_Тромбоциты_PLT": 124,
56     "ОАК_ШиринаРаспредТромбоцитов_PDW": "18,5",
57     "ОАК_Эозинофилы_EOS": null,
58     "ОАК_Эозинофилы_EOS%": null,
59     "ОАК_Эритроциты_RBC": "5,07",
60     "ОАМ_Глюкоза": "В норме",
61     "ОАМ_Кетоны": "0 Отрицательно",
62     "ОАМ_Кислотность": "Кислая",
63     "ОАМ_Лейкоциты": "2,5",
64     "ОАМ_Прозрачность": "Прозрачная",
65     "ОАМ_УдельныйВес": "1,026",
66     "ОАМ_Цвет": "Желтый",
67     "ОАМ_ЭпителийПлоский": "3,4",
68     "ОАМ_Эритроциты": "5,50",
69     "ПротромбиновоеВремя": "24,7",
70     "ПротромбиновыйИндекс": "35,0",
71     "Рост": "173",
72     "Триглицериды": "0,89",
73     "Фибриноген": "270,0",
74     "Холестерин": "3,02",
75     "ОАК_Ядерные_эритроциты_NRBC": null,
76     "ОАК_Ядерные_эритроциты_NRBC%": null,
77     "АортальныйСтеноз": null,
78     "КДР": null,
79     "КСР": null,
80     "МитральнаяРегургитация": null,
81     "ОбъемЛП": null,
82     "СДЛА": null,
83     "ТрикуспидальнаяРегургитация": null,
84     "ТрикуспидальныйСтеноз": null,
85     "ФBSimpson": null,
86     "ФВTeichgoIz": null,
87     "Возраст": 55.0
88 }'

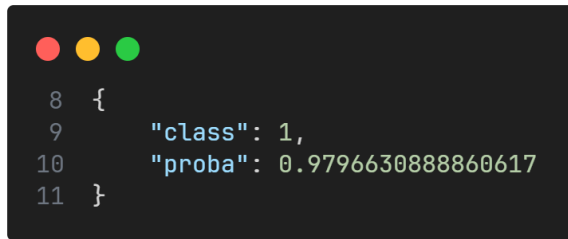
```

Рисунок 18 - Результат запроса получения списка признаков

Тестирование метода получения оценки риска ухудшения состояния здоровья «predictionOfHealthDeterioration/api/SZZ/predict» представлено на рисунках 19 и 20.

```
8 curl --location 'http://127.0.0.1:5000/api/predict' \  
9 --header 'Content-Type: application/json' \  
10 --data '{  
11   "Пол": "М",  
12   "Диагноз": "I50.0",  
13   "CRP": null,  
14   "DIA": 40.0,  
15   "NT_proBNP_экспресс": "22 037,00",  
16   "SIS": 80.0,  
17   "БАК_АЛТ": "99,5",  
18   "БАК_АСТ": "104,2",  
19   "БАК_АЧТВ": "33,1",  
20   "БАК_БелокОбщий": "58,3",  
21   "БАК_БилирубинОбщий": "30,5",  
22   "БАК_БилирубинПрямой": "17,70",  
23   "БАК_Калий": "5,20",  
24   "БАК_Креатинин": "246,0",  
25   "БАК_ЛипопротеидыВысокойПлотности": "0,67",  
26   "БАК_ЛипопротеидыНизкойПлотности": "2,34",  
27   "БАК_СкоростьКлубочковойФильтрацииMDRD": "25,35",  
28   "БАК_ТропонинI": "0,014",  
29   "Вес": "126.",  
30   "КФК": "150,0",  
31   "ЛДГ": null,  
32   "МВ_КФК": "20,3",  
33   "МНО": "2,24",  
34   "ОАК_Базофилы_BASO": "0,010",  
35   "ОАК_Базофилы_BASO%": "0,1",  
36   "ОАК_Гематокрит_HCT": "50,9",  
37   "ОАК_Гемоглобин_HGB": 154.0,  
38   "ОАК_Глюкоза": "5,88",  
39   "ОАК_КоэффициентЭритроц_РDWCV": "14,1",  
40   "ОАК_Лейкоциты_WBC": "11,29",  
41   "ОАК_Лимфоциты_LYM": "2,27",  
42   "ОАК_Лимфоциты_LYM%": "20,1",  
43   "ОАК_Моноциты_MONO": "0,44",  
44   "ОАК_Моноциты_MONO%": "3,9",  
45   "ОАК_Мочевина": "20,00",  
46   "ОАК_НезрелыеГранулоциты_IG": "0,02",  
47   "ОАК_НезрелыеГранулоциты_IG%": "0,20",  
48   "ОАК_Нейтрофилы_NEU": "8,57",  
49   "ОАК_Нейтрофилы_NEU%": "75,9",  
50   "ОАК_Сред_конц_гемоглобина_в_эр_МСНС": 337.0,  
51   "ОАК_Сред_сод_гемоглобина_эритроците_МСН": "30,4",  
52   "ОАК_Средний_объем_эритроците_MCV": "90,1",  
53   "ОАК_СреднийОбъемТромбоцитов_MPV": "12,9",  
54   "ОАК_Тромбоциты_PCT": "0,2",  
55   "ОАК_Тромбоциты_PLT": 124,  
56   "ОАК_ШиринаРаспредТромбоцитов_PDW": "18,5",  
57   "ОАК_Эозинофилы_EOS": null,  
58   "ОАК_Эозинофилы_EOS%": null,  
59   "ОАК_Эритроциты_RBC": "5,07",  
60   "ОАМ_Глюкоза": "В норме",  
61   "ОАМ_Кетоны": "0 Отрицательно",  
62   "ОАМ_Кислотность": "Кислая",  
63   "ОАМ_Лейкоциты": "2,5",  
64   "ОАМ_Прозрачность": "Прозрачная",  
65   "ОАМ_УдельныйВес": "1,026",  
66   "ОАМ_Цвет": "Желтый",  
67   "ОАМ_ЭпителийПлоский": "3,4",  
68   "ОАМ_Эритроциты": "5,50",  
69   "ПротромбиновоеВремя": "24,7",  
70   "ПротромбиновыйИндекс": "35,0",  
71   "Рост": "173",  
72   "Триглицериды": "0,89",  
73   "Фибриноген": "270,0",  
74   "Холестерин": "3,02",  
75   "ОАК_Ядерные_эритроциты_NRBC": null,  
76   "ОАК_Ядерные_эритроциты_NRBC%": null,  
77   "АортальныйСтеноз": null,  
78   "КДР": null,  
79   "КСР": null,  
80   "МитральнаяРегургитация": null,  
81   "ОбъемПП": null,  
82   "СДЛА": null,  
83   "ТрикуспидальнаяРегургитация": null,  
84   "ТрикуспидальныйСтеноз": null,  
85   "ФBSimpson": null,  
86   "ФВTeichholz": null,  
87   "Возраст": 55.0  
88 }'
```

Рисунок 19 - Запрос получения оценки риска ухудшения состояния здоровья

A terminal window with a dark background and three colored window control buttons (red, yellow, green) at the top left. It displays a JSON object with two fields: "class" and "proba".

```
8 {  
9     "class": 1,  
10    "proba": 0.9796630888860617  
11 }
```

Рисунок 20 - Результат запроса получения оценки риска ухудшения состояния здоровья

3.3 Вывод

Реализован веб-сервис для взаимодействия МИС с обученной моделью на базе языка программирования Python при использовании библиотек Pickle и Flask. Описаны используемые методы, проверена работоспособность каждого.

4 РАЗРАБОТКА ПОДСИСТЕМЫ МИС «1С:МЕДИЦИНА.РЕГИОН»

МИС «1С:Медицина.Регион» разработана на платформе «1С:Предприятие 8», на основе лежит типовая конфигурация «1С:Медицина.Больница». Система предоставляет обширные возможности медицинскому персоналу. В ней ведется учет оказанных услуг в разрезе клинических отделений и медицинских рабочих мест, регистрируется назначение лекарств, исследований и т.д.

Сотрудники МО, в зависимости от своей должности, пользуются теми или иными возможностями и функциями МИС для выполнения своей работы и обработки медицинской информацией. Медицинская информация - информация, относящаяся непосредственно к человеку как пациенту, то есть информация о его здоровье, особенностях организма, перенесенных заболеваниях и др. Использование МИС при работе с медицинской информацией подразумевает отсутствие необходимости куда-либо идти для ее получения, взаимодействуя со стационарным компьютером – всё хранится в базе данных на сервере.

4.1 Проектирование

Сначала были определены основные функции, которые должна выполнять подсистема для представления пользователю о прогнозируемых случаях ухудшения состояния здоровья:

- регистрация и формирование списка показателей здоровья для получения оценки после получения результатов лабораторных исследований в случае, если хотя бы по одному из целевых видов показателей здоровья пришли новые данные;
- получение прогноза на основании зарегистрированного списка показателей здоровья;
- отображение информации о возможном риске неблагоприятного исхода и уведомление пользователя о таком случае после получения оценки;

- учет всех отправленных пакетов на прогнозирование и их результатов в разрезах: время, пациент, диагноз.
- получение списка видов показателей здоровья из веб-сервиса, на которых обученной моделью проводится прогнозирование.

На основании выявленных функциональных требований были определены основные объекты конфигурации расширения и их структура:

- справочники:
 - «Соответствия видов показателей здоровья и диагнозов»;
 - «Картотека» (заимствованный из основной конфигурации);
 - «МКБ-10» (заимствованный из основной конфигурации);
 - «Значения показателей здоровья» (заимствованный из основной конфигурации);
- документы:
 - «Регистрация случаев прогнозирования ухудшения состояния здоровья»;
 - «Регистрация результатов прогнозирования».
- регистры сведений:
 - «Показатели здоровья случаев прогнозирования»;
 - «результаты прогнозирования».
- Обработки:
 - «Настройки подключения к веб-сервису модели»;
 - «АРМ подсистемы прогнозирования».
- Планы видов характеристик:
 - «Виды показателей здоровья» (заимствованный из основной конфигурации);
- Перечисления:
 - «Оценки результатов прогнозирования»;
 - «HL7v3 type» (заимствованный из основной конфигурации).

Таблица 3 - Структура реквизитов справочника "Соответствия диагнозов и видов показателей здоровья

Реквизит	Тип данных
НозологическаяГруппа	СправочникСсылка.тмб_НозологическиеГруппы
Код	Число
Наименование	Строка
ПометкаУдаления	Булево
ВидыПоказателейЗдоровья	ТабличнаяЧасть

Таблица 4 - Структура табличной части "ВидыПоказателейЗдоровья" справочника "Соответствия диагнозов и видов показателей здоровья

Реквизит	Тип данных
ВидПоказателяЗдоровья	ПланВидовХарактеристикСсылка.ВидыПоказателейЗдоровья
Соответствие	Строка

Таблица 5 - Структура реквизитов справочника "Картотека"

Реквизит	Тип данных
ИдентификаторПациентаВРМИС	Число
СтраховойНомерПФР	Строка
КодВТФОМС	Число
тмб_ИНН	Строка
ДатаВыбытия	Дата
ДокументВыбытия	ОпределяемыйТип.ДокументВыбытияПациента
сфпПользовательДляПереключенияЗвонков	СправочникСсылка.Пользователи

тмб_ИдентификаторАнонимногоПациента	Число
Код	Число
Наименование	Строка
ПометкаУдаления	Булево

Таблица 6 - Структура реквизитов справочника "МКБ10"

Реквизит	Тип данных
Код	Строка
Наименование	Строка
ПометкаУдаления	Строка
КодМинздрава	Строка
НаименованиеПолное	Строка
OIDКлассификатораМинздрава	Строка
ВерсияКлассификатораМинздрава	Строка
ДополнительныйКод	Строка
ИдентификаторНаДиске	Число
ИспользуетсяВОМС	Булево
Актуальность	Булево
ВАрхиве	Булево

Таблица 7 - Структура реквизитов справочника "ЗначенияПоказателейЗдоровья"

Код	Число
Наименование	Строка
ПометкаУдаления	Булево
АтрибутCode	Строка
Идентификатор	Строка
codeSystemName	Строка

codeSystem	Строка
ФлагПоказателяЗдоровья	ПеречислениеСсылка.ФлагиПоказателейЗдоровья
НаименованиеПолное	Строка

Таблица 8 - Структура документа "Регистрация случаев прогнозирования ухудшения состояния пациента"

Реквизит	Тип данных
Код	Число
Проведен	Булево
ПометкаУдаления	Булево
Пациент	СправочникСсылка.Картотека
Диагноз	СправочникСсылка.МКБ10
ПоказателиЗдоровья	ТабличнаяЧасть

Таблица 9 - Структура табличной части "ПоказателиЗдоровья" документа "Регистрация случаев прогнозирования ухудшения состояния здоровья пациента"

Реквизит	Тип данных
Дата	Дата
ВидПоказателяЗдоровья	ПланВидовХарактеристикСсылка.ВидыПоказателейЗдоровья
Значение	ХарактеристикаСсылка.ВидыПоказателейЗдоровья
ТипЗначения	ПеречислениеСсылка.NL7v3Type

Таблица 10 - Структура реквизитов документа "Регистрация результатов прогнозирования"

Реквизит	Тип данных
ДокументОснования	ДокументСсылка.РегистрацияСлучаевПрогнозирования

Оценка	ПеречислениеСсылка. ПУСЗП_ОценкиРезультатовПрогнозирования
Вероятность	Число
СтатусКод	Число
СтатусТекст	Строка

Таблица 11 - Структура реквизитов регистра сведений "Показатели здоровья прогнозирования"

Реквизит	Тип данных
Пациент	СправочникСсылка.Картотека
Диагноз	СправочникСсылка.Диагноз
ВидПоказателяЗдоровья	ПланВидовХарактеристикСсылка.ВидыПоказателей Здоровья
Значение	ХарактеристикаСсылка.ВидыПоказателейЗдоровья
Дата	Дата
ТипЗначения	ПеречислениеСсылка.NL7v3Type

Таблица 12 - Структура реквизитов регистра сведений "Результаты прогнозирования"

Реквизит	Тип данных
ДокументОснование	ДокументСсылка.РегистрацияСлучаевПрогнозирования
Оценка	ПеречислениеСсылка. ПУСЗП_ОценкиРезультатовПрогнозирования
Вероятность	Число
СтатусКод	Число
СтатусТекст	Строка
СотрудникОзнакомлен	Булево

4.2 Реализация

Для разработки подсистемы было решено реализовывать расширение в виду следующих положительных моментов:

- оперативное внесение изменений и исправление ошибок;
- отсутствие зависимости от обновления конфигурации 1С;
- сохранение поддержки объектов конфигурации;
- отключение подсистемы в непредвиденных случаях.

Расширение является одним из инструментов быстрой разработки, позволяя вносить доработки в основную конфигурацию без ее непосредственного изменения и обновления [15].

В среде разработки «1С:Предприятие 8» было создано расширение «Подсистема прогнозирования ухудшения состояния здоровья пациентов» с типом назначения «Дополнение», область действия – «Информационная база», для новых объектов расширения задан префикс «ПУСЗП_» (см. рисунок 21). Все добавленные объекты в конфигурацию расширения будут иметь префикс «ПУСЗП».

Свойства	
Поиск (Ctrl+Alt+l)	
▼ Основные:	
Имя	ПодсистемаПрогнозированияУхудшенияСостоянияЗдоровьяПациентов
Синоним	Подсистема прогнозирования ухудшения состояния здоровья пациентов
Комментарий	
Назначение расширения конфигурации	Дополнение
Поддерживать соответствие объектам расширяемой конфиг	<input checked="" type="checkbox"/>
Префикс имен	ПУСЗП_
Принадлежность объекта	Заимствованный
<input type="checkbox"/> Объект расширяемой конфигурации	
<input checked="" type="checkbox"/> Основной режим запуска	Управляемое приложение
<input checked="" type="checkbox"/> Назначения использования	Приложение для платформы
Вариант встроенного языка	Русский
<input checked="" type="checkbox"/> Основные роли	Роль.ПУСЗП_ОсновнаяРоль

Рисунок 21 - Свойства разрабатываемого расширения

Справочники.

Справочники используются для хранения редко-изменяемых данных и последующего использования элементов в сторонних объектах. В текущем случае, был разработан справочник хранения соответствий «нозологических групп» и видов показателей здоровья (см рисунок 22).

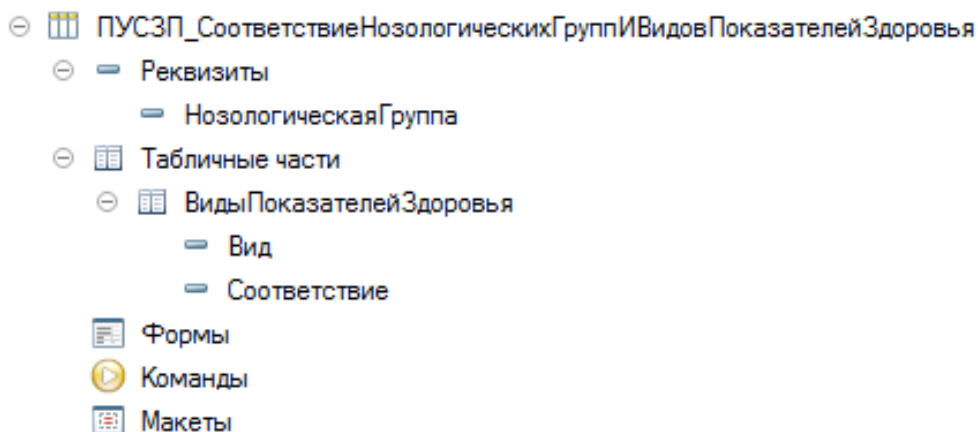


Рисунок 22 – Разработанный справочник соответствий нозологических групп и видов показателей здоровья

«Нозологическая группа» представляет из себя группу диагнозов, представляющих определенное заболевание. Справочник данных значений был уже реализован в рамках задачи «ведения нозологических регистров», и полностью подходит для использования в текущей разработке.

Из основной конфигурации были заимствованы справочники, используемые в программном коде, запросах и как тип данных в других объектах (см рисунок 23).

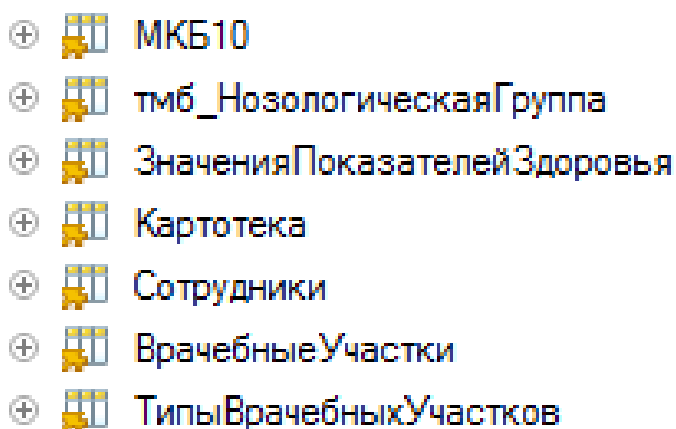


Рисунок 23 - Заимствованные из основной конфигурации справочники Перечисления.

Перечисления в «1С» используются для хранения и использования неизменяющихся пользователями predetermined значений. Ведением перечислений обычно занимаются разработчики.

В расширении было создано перечисление с оценками риска ухудшения состояния здоровья с двумя значениями – «ЕстьРиск» и «НетРиска» (см. рисунок 24).

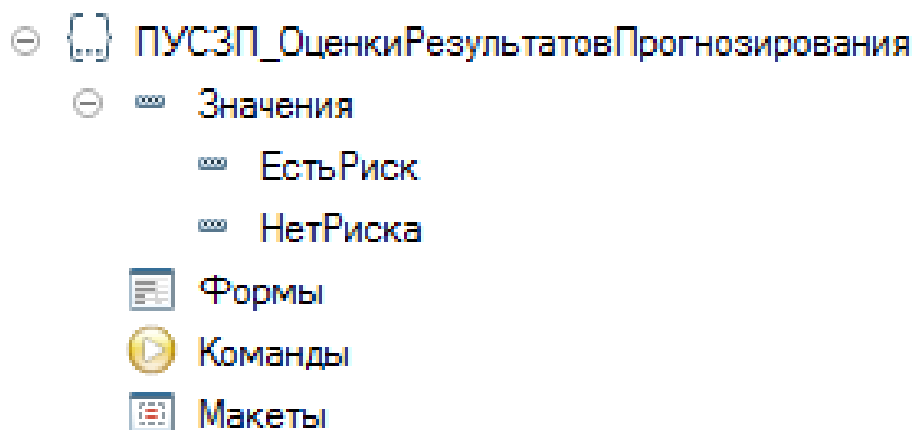


Рисунок 24 – Созданное перечисление оценки риска ухудшения состояния здоровья

Из основной конфигурации было заимствовано перечисление типов значений видов показателей здоровья, используемое в документах и регистрах сведений (см. рисунок 25).



Рисунок 25 - Заимствованное из основной конфигурации перечисление Планы видов характеристик

Из основной конфигурации были заимствованы «Виды показателей здоровья», которые представлены в виде объекта плана видов характеристик, который предназначен для хранения различных свойств или характеристик описываемых объектов, которые задаются пользователями для дальнейшего использования (см. рисунок 26).

Виды Показателей Здоровья

Рисунок 26 - Заимствованные из основной конфигурации планы видов характеристик

Документы.

В расширении были разработаны документы регистрации случаев прогнозирования и регистрации оценки прогнозирования, предназначенные для ведения контроля прогнозируемых данных и получаемых из веб-сервиса оценок соответственно (см. рисунок 27). На прямую с данными документами пользователь не работает, они используются как служебные объекты.

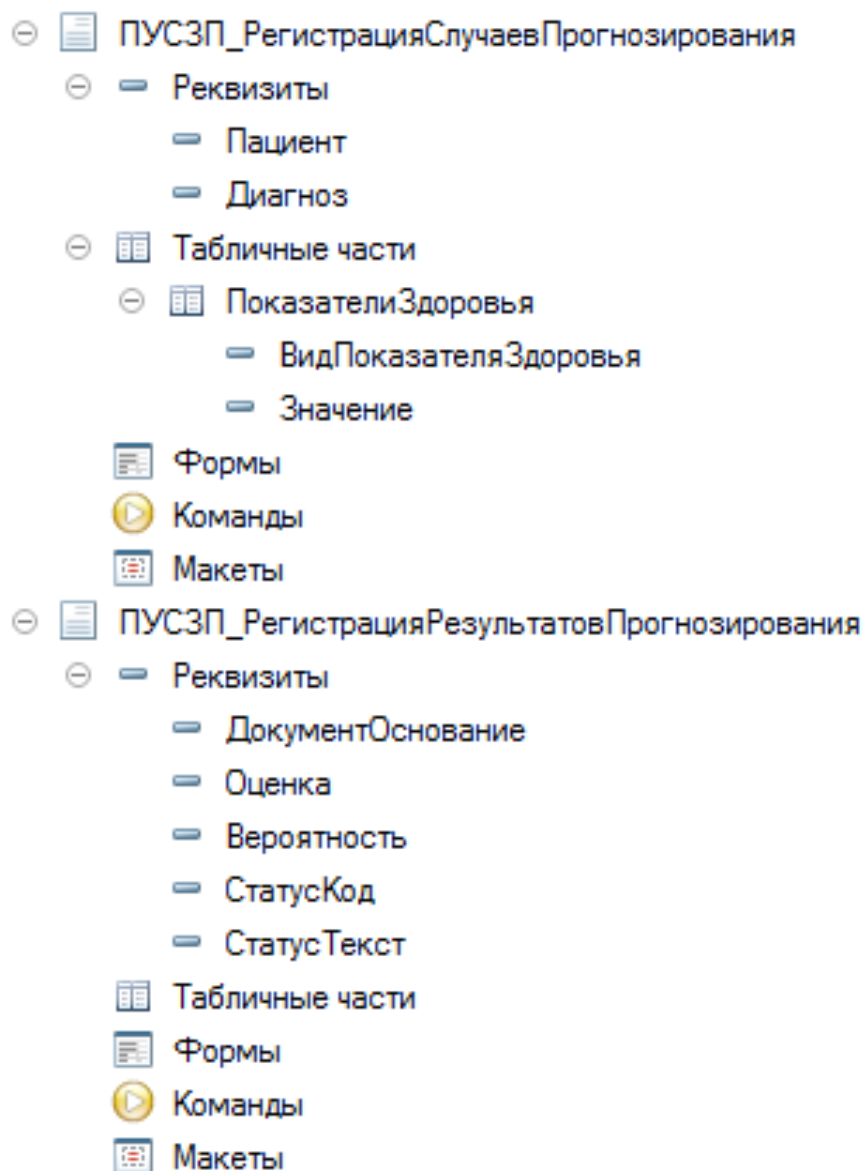


Рисунок 27 – Разработанные в расширении документы

Регистры сведений.

В дополнение к документам для хранения прогнозируемых случаев и оценки в разрезах времени, регистраторов, пациентов и диагнозов были разработаны регистры сведений (см. рисунок 28), которые в последствии будут использоваться для составления отчетов, печатных форм, а также контроля работы подсистемы.

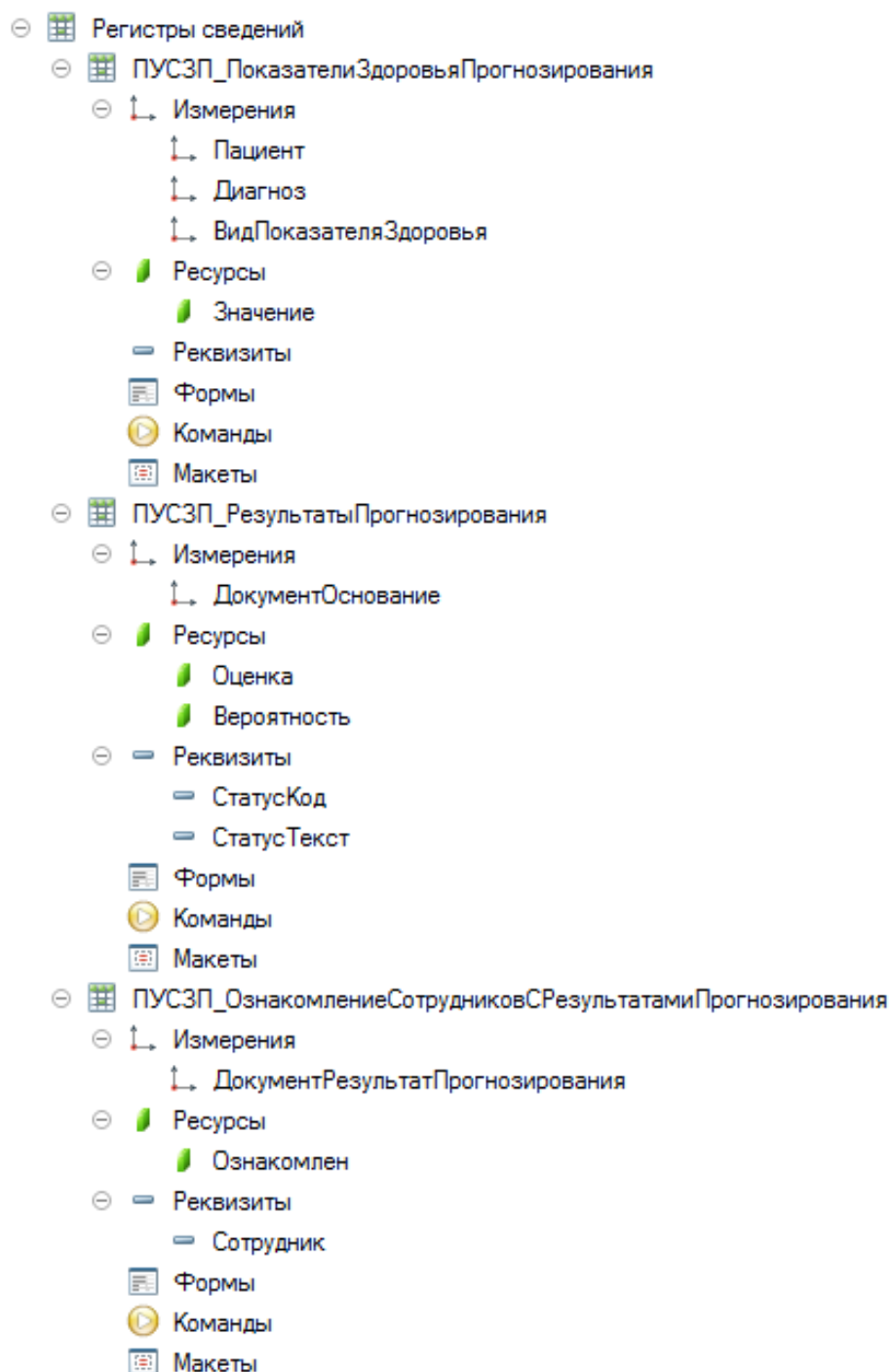


Рисунок 28 – Разработанные в расширении регистры сведений

Обработки.

В расширении были разработаны две обработки (см. рисунок 29):

- для настройки подключения к веб-сервису;
- просмотра списка «текущих дел» пользователя.

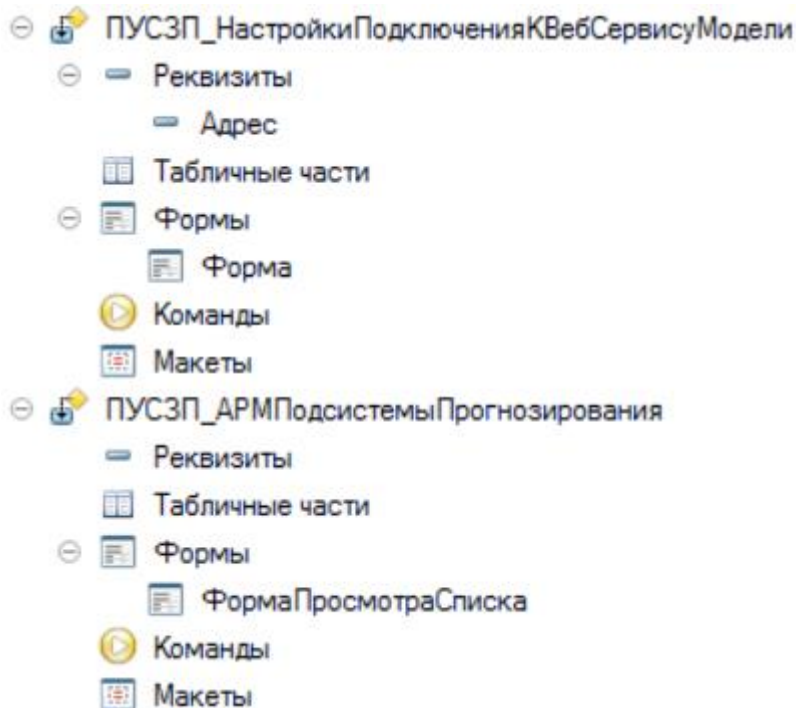


Рисунок 29 - Разработанные в расширении обработки

На текущий момент, аутентификация в веб-сервисе не ведется, так как сам веб-сервис проходит тестирование и опубликован на закрытом локальном контуре сети организации «1С-Медицина-Регион».

Общие модули.

Весь общий и часто используемый код был вынесен в соответствующие разработанные общие модули (см. рисунок 30).

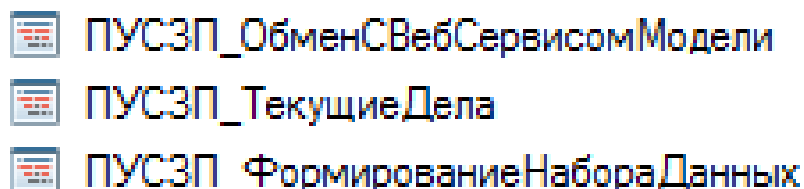


Рисунок 30 - Разработанные в расширении общие модули

Из основной конфигурации были заимствованы следующие общие модули для переопределения определенных методов (см. рисунок 31):

- Тбит_ВебСервисЛИС – переопределение метода регистрация результатов лабораторных исследований;
- ТекущиеДелаПереопределяемый – переопределение метода формирования текущих дел пользователя.

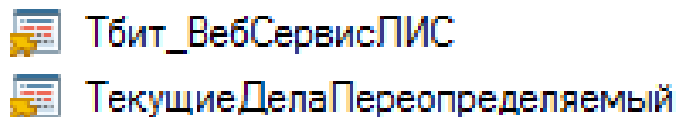


Рисунок 31 - Заимствованные из основной конфигурации общие модули

Одним из существующих инструментов уведомления пользователя о различных случаях оперативного вмешательства – «Текущие дела». Данный механизм представляет из себя автоматически обновляющийся список с разными категориями «дел», каждый из которых должен быть обработан пользователем вручную. Доступ к «текущим делам» есть как с начальной страницы приложения (см. рисунок 32), так и со страницы «Текущие дела» объекта «АРМ врача» подсистемы «Контроль исполнения» (см. рисунок 33).

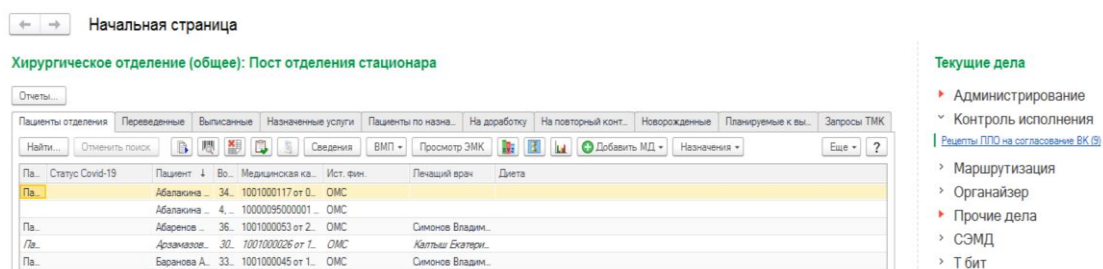


Рисунок 32 – инструмент «Текущие дела» на начальной странице

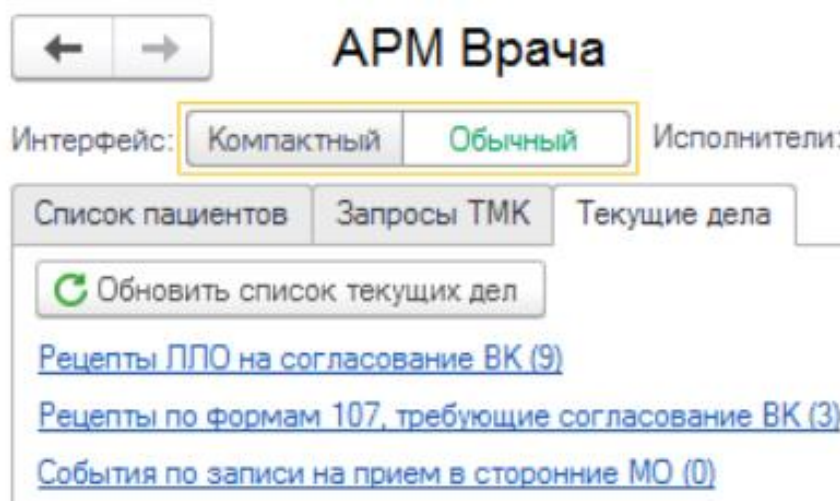


Рисунок 33 – инструмент «Текущие дела» на форме объекта «АРМ врача»

Настройка подсистемы начинается с заполнения адреса веб-сервиса и заполнения справочника соответствий нозологических групп и видов показателей здоровья (см. рисунок 34 и 35).

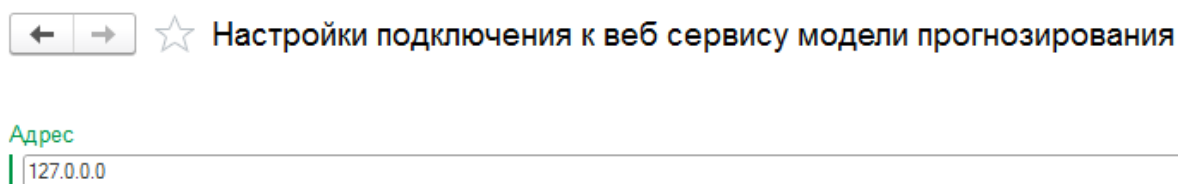


Рисунок 34 - Заполнение адреса веб-сервиса

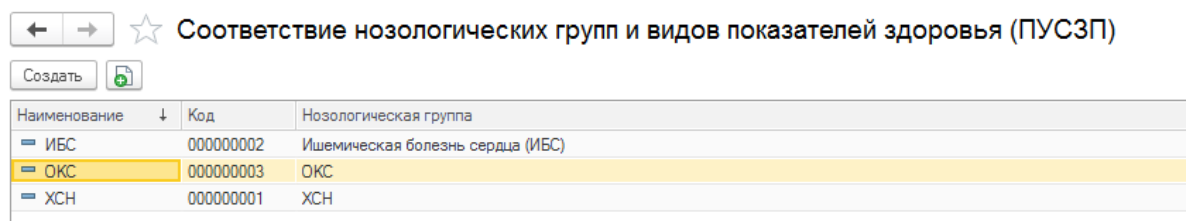


Рисунок 35 - Создание элементов справочника соответствий нозологических групп и видов показателей здоровья

Пример заполнения элемента справочника соответствий нозологических групп и видов показателей здоровья представлен на рисунке 36.

☆ ОКС (Соответствие нозологических групп и видов показателей здоровья ...)

Записать и закрыть Записать Еще ▾

Код: 000000003

Наименование: ОКС

Нозологическая группа: ОКС

Добавить ↑ ↓ Еще ▾

N	Вид	Соответствие
66	Эритроциты	OAM_Эритроциты
67	Митральная регургитация	МитральнаяРегургитация
68	TP	БАК_БелокОбщий
69	Моноциты (MONO%)	OAK_Моноциты_MONO%
70	Вес	Вес
71	CHOL	Холестерин
72	NT-проBNP (N-проМНП, экспресс)	NT_проBNP_Экспресс
73	Рост	Рост
74	K (Калий)	БАК_Калий
75	NEU% (Нейтрофилы,%)	OAK_Нейтрофилы_NEU%
76	Тромбоциты	OAK_Тромбоциты_PLT
77	BAS# (Базофилы,#)	OAK_Базофилы_BASO
78	Лейкоциты (полуколич.)	OAM_Кислотность
79	Базофилы,% (микроскопия)	OAK_Базофилы_BASO%
80	NEU# (Нейтрофилы,#)	OAK_Нейтрофилы_NEU
81	Прозрачность	OAM_Кетоны
82	MON% (Моноциты,%)	OAK_Моноциты_MONO%
83	Гематокрит(НСТ)	OAK_Гематокрит_НСТ
84	AST (Аспаратаминотрансфераза)	БАК_ACT
85	MCH (Сред. содер. HGB в RBC)	OAK_Сред_сод_гемоглобина_эритроците_MCH
86	Прозрачность	OAM_Прозрачность
87	Лейкоциты	OAM_Лейкоциты
88	АЧТВ	БАК_АЧТВ
89	Креатинин	БАК_Креатинин
90	TBIL	БАК_БилирубинОбщий
91	HDL (ЛПВП)	БАК_ЛипопротеидыВысокойПлотности
92	Группа крови	ГруппаКрови
93	Цвет	OAM_Цвет
94	RBC# (Эритроциты,#)	OAK_Эритроциты_RBC
95	Dia	DIA
96	Вес	Вес
97	Общий белок	БАК_БелокОбщий
98	PDW-SD (Шир. распр. PLT по V)	OAK_ШиринаРаспредТромбоцитов_PDW
99	MCV (Сред. объем эритроцита)	OAK_Средний_объем_эритроцита_MCV
100	INR (МНО)	МНО
101	LYM# (Лимфоциты,#)	OAK_Лимфоциты_LYM
102	Лейкоциты(WBC)	OAK_Лейкоциты_WBC
103	WBC# (Лейкоциты,#)	OAK_Лейкоциты_WBC
104	Уробилиноген	OAM_ЭпителийПлоский

Рисунок 36 - Пример заполнения элемента справочника соответствий нозологических групп и видов показателей здоровья

При включенном расширении каждая регистрация видов показателей здоровья, а именно создание соответствующего медицинского документа «Регистрация лабораторных исследований» (см. рисунок 37) начинает процесс прогнозирования риска ухудшения состояния.

Анализ крови биохимический для МИГ

Пациент: Т... ч **№ карты:** 0101012033
Дата рождения: 25.01.1963 **Пол:** Мужской
Местожительство:
Назначил: **Дата выполнения:** 02.06.2023
Исп. врач: врач-пульмонолог Ко... на **Время выполнения:** 15:02
Исп. МО: ГБУЗ ТО "ОКБ №1" **Телефон:** +7(3452)287400
Адрес: 625023, Тюменская обл, Тюмень г, Котовского ул, дом № 55

Наименование	Результат	Ед.изм.	Норма	Дата
Общий белок	72,1	Г/Л	64,0 - 83,0	02.06.2023 15:02:59
Мочевина	4,6	ММОЛЬ/ Л	2,8 - 8,1	02.06.2023 15:02:59
Креатинин	68	МКМОЛЬ /Л	62,0 - 106,0	02.06.2023 15:02:59
Билирубин общий	14	МКМОЛЬ /Л	0,0 - 21,0	02.06.2023 15:02:59
Билирубин прямой	1,7	МКМОЛЬ /Л	0,0 - 5,0	02.06.2023 15:02:59
АСТ	31,2	ЕД/Л	0,0 - 40,0	02.06.2023 15:02:59
АЛТ	12,1	ЕД/Л	0,0 - 41,0	02.06.2023 15:02:59

Рисунок 37 - Пример медицинского документа " Регистрация лабораторных исследований"

После регистрации результатов лабораторных исследований проверяется наличие диагноза из нозологических групп, которые были заполнены в соответствующем справочнике. В случае соответствия диагноза пациента и нозологической группы, создается документ «Регистрация случаев прогнозирования ухудшения состояния здоровья» (см. рисунок 38) с заполнением рассматриваемых видов показателей здоровья и их значениями (см рисунок 39).

← → ☆ Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП)

Создать

Дата	Номер	Пациент	Диагноз
01.06.2023 17:27:51	000000008	Б...	I22.1
02.06.2023 12:14:48	000000004	И...	I21.9
02.06.2023 15:03:02	000000006	Т...	I25.1
05.06.2023 4:00:49	000000002	С...	I22.9
06.06.2023 10:00:15	000000003	Е...	I25.1
06.06.2023 14:45:08	000000005	К...	I20.0
07.06.2023 8:52:56	000000009	Т...	I25.9
10.06.2023 23:17:26	000000001	К...	I21.1
14.06.2023 10:32:17	000000010	И...	I20.8
15.06.2023 10:08:06	000000007	Т...	I25.9

Рисунок 38 - Форма списка документа "Регистрация случаев прогнозирования ухудшения состояния здоровья"

← → ☆ Регистрация случаев прогнозирования

Провести и закрыть Записать Провести

Номер: 000000006

Дата: 02.06.2023 15:03:02

Пациент: Т...

Диагноз: I25.1

Добавить

N	Вид показателя здоровья	Значение
18	Нитриты	1 012
19	Лейкоциты(WBC)	8,96
20	Трикуспидальная регургитация	
21	Ядерные эритроциты (NRBC)	
22	Базофилы, % (микроскопия)	0,3
23	Эозинофилы, % (микроскопия)	1
24	Dia	90
25	LYM# (Лимфоциты, #)	2,3
26	K (Калий)	4,7
27	Уробилиноген	0,0 единиц
28	Белок в моче (колич.)	0,00 0-1
29	LDL (ЛПНП)	
30	Резус-фенотипирование	
31	MPV (Сред. объем тромбоцита)	9,4
32	Лимфоциты, % (микроскопия)	25,7
33	PTT (Протромбиновое время)	12,1
34	Аортальный стеноз	
35	Фибриноген	3,88
36	Вес	112
37	Моноциты, % (микроскопия)	6,8
38	PLT# (Тромбоциты, #)	213
39	Сред. сод. гемоглобина эритроците (MCH)	29,8
40	INR (MHO)	1,14
41	Кислотно-щелочная реакция	
42	Кетоны	Прозрачная
43	АЧТВ	24,5
44	СК (Креатинкиназа)	132
45	Трикуспидальный стеноз	
46	UREA (Мочевина)	4,6
47	MON# (Моноциты, #)	0,61
48	TBIL	14
49	RDW-CV (Шир. распр. RBC по V)	17,1

Рисунок 39 - Заполнение показателей здоровья документа При проведении документа регистрации случаев прогнозирования выполняется следующий процесс:

- запрос «/predictionOfHealthDeterioration/api/SSZ/getFeatures» для получения списка соответствий видов показателей здоровья, на которых обучалась модель;
- формирование набора данных пациента на основании полученного списка признаков;
- запрос «/predictionOfHealthDeterioration/api/SSZ/predict» для получения оценки риска ухудшения состояния;
- создание и проведение документа «Регистрация результатов прогнозирования ухудшения состояния здоровья» (см. рисунок 40), в котором заполняются реквизиты «Документ-основание», «Оценка», «Вероятность», «Статус код», «Статус текст».

← → ☆ Регистрация результатов прогнозирования ухудшения состояния здоровья (ПУСЗП)

Создать

Дата	↓	Номер	Документ основание	Оценка	Вероятность	Статус код	Статус текст
01.06.2023 17:27:51		00000...	Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) ...	Есть риск	98,122	200	
02.06.2023 12:14:50		00000...	Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) ...	Есть риск	78,920	200	
02.06.2023 15:03:04		00000...	Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) ...	Есть риск	89,022	200	

Рисунок 40 - Форма списка документа "Регистрация результатов прогнозирования ухудшения состояния здоровья"

Проведение документов «Регистрации случаев прогнозирования» и «Регистрации результатов прогнозирования» вносят данные в соответствующие регистры сведений, представленные на рисунках 41 и 42.

← → ☆ Показатели здоровья прогнозирования (ПУСЗП)

Регистратор	Номер	П.	Диагноз	Вид показателя здоровья	Значение
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	21	H.	I21.9	PTI (Протромбиновый индекс)	88
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	48	H.	I21.9	Нитриты	1 015
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	37	H.	I21.9	NT-proBNP (N-проМНП, экспресс)	750,4
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	45	H.	I21.9	RDW-CV (Шир. распр. RBC по V)	14,7
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	68	H.	I21.9	PDW-SD (Шир. распр. PLT по V)	9,9
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	14	H.	I21.9	MPV (Сред. объем тромбоцита)	9,5
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	29	H.	I21.9	PCT (Тромбокрит)	0,1
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	49	H.	I21.9	NEU# (Нейтрофилы, #)	7,31
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	16	H.	I21.9	LYM# (Лимфоциты, #)	2
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	28	H.	I21.9	MON# (Моноциты, #)	0,63
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	57	H.	I21.9	BAS# (Базофилы, #)	0,05
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	23	H.	I21.9	NEU% (Нейтрофилы, %)	71
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	65	H.	I21.9	UREA (Мочевина)	2,9
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	3	H.	I21.9	DBIL (Билирубин прямой)	
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	59	H.	I21.9	LDL (ЛПНП)	2,04
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	6	H.	I21.9	HDL (ЛПВП)	1,17
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	47	H.	I21.9	INR (МНО)	1,08
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	22	H.	I21.9	TG (Триглицериды)	1,21
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	54	H.	I21.9	TBIL	
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	9	H.	I21.9	AST	30,4
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	56	H.	I21.9	Фибриноген	710
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	30	H.	I21.9	Лейкоциты(WBC)	10,3
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	44	H.	I21.9	Эритроциты(RBC)	3,3
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	15	H.	I21.9	Гемоглобин(HGB)	99
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	61	H.	I21.9	Средний объем эритроцита(MCV)	87,9
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	63	H.	I21.9	Сред. сод. гемоглобина эритроците (MCH)	30
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	64	H.	I21.9	Сред. конц. гемоглобина в эр. (MCHC)	341
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	8	H.	I21.9	АЛТ	26,7
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	52	H.	I21.9	Реус-фенотипирование	ccDEe
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	10	H.	I21.9	С-реактивный белок (качественный)	6,812
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	67	H.	I21.9	TnI (ТропонинI, экспресс-тест)	3 337
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	34	H.	I21.9	Группа крови	A?B(V) че
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	50	H.	I21.9	Реус-принадлежность	Rh(+) Поло
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	40	H.	I21.9	Общий белок	62,1
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	51	H.	I21.9	Митральная регургитация	Per. (++)
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	17	H.	I21.9	Трикуспидальная регургитация	Per. (min)
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	38	H.	I21.9	PTT (Протромбиновое время)	14,1
Регистрация случаев прогнозирования ухудшения состояния здоровья (ПУСЗП) 000000004 от 02.06.2023	46	H.	I21.9	Креатинин	68

Рисунок 41 - Пример формы списка регистра сведений "Показатели здоровья прогнозирования"

← → ☆ Результаты прогнозирования (ПУСЗП)

Период	Регистратор	Номер с.	Документ основание	Оценка	Вероятность	Статус код	Статус текст
01.06.2023 17:27:51	Регистрация результатов прогнозирования...	1	Регистрация случаев прогнозирования ...	Есть риск	98,122	200	
02.06.2023 12:14:50	Регистрация результатов прогнозирования...	1	Регистрация случаев прогнозирования ...	Есть риск	78,920	200	
02.06.2023 15:03:04	Регистрация результатов прогнозирования...	1	Регистрация случаев прогнозирования ...	Есть риск	89,022	200	
20.06.2023 19:06:34	Регистрация результатов прогнозирования...	1	Регистрация случаев прогнозирования ...	Нет риска	80,033	200	
20.06.2023 19:07:43	Регистрация результатов прогнозирования...	1	Регистрация случаев прогнозирования ...	Есть риск	92,100	200	

Рисунок 42 - Пример формы списка регистра сведений "Результаты прогнозирования"

Работа пользователя с подсистемой начинается при открытии МИС «1С:Медицина.Регион», на начальной странице которой, в случае наличия нерассмотренных результатов прогнозирования, формируется команда на вкладке текущих дел (см. рисунок 43).

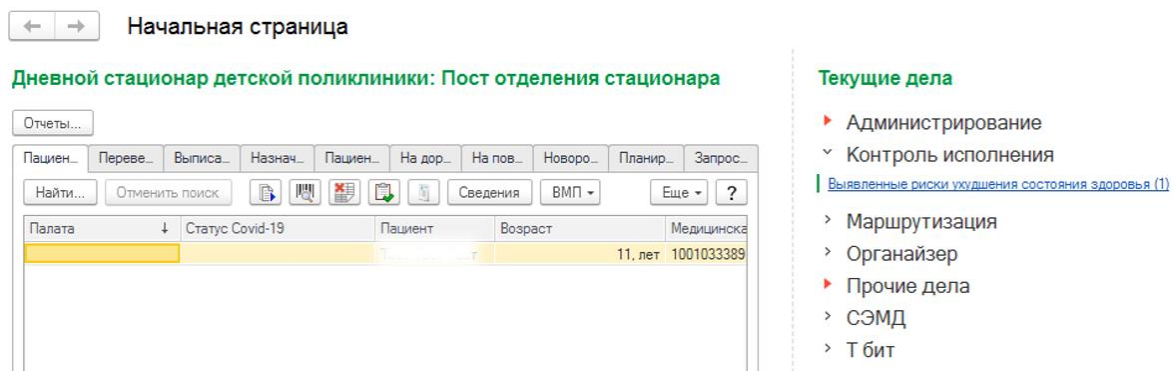


Рисунок 43 – Вкладка «Текущих дел» на начальной странице

При нажатии по гиперссылке «Выявленные риски ухудшения состояния здоровья» открывается форма списка нерассмотренных результатов прогнозирования, для которых была получена оценка «ЕстьРиск» (см рисунок 44).

При выделении строки подгружаются виды показателей здоровья, на которых была спрогнозирована оценка риска. На форме дополнительно отображаются команды «Ознакомился» и «Открыть ОДП».

Команда «Ознакомился» предназначена для контроля учета ознакомления сотрудника с выявленным риском.

Команда «Открыть ОДП» предназначена для открытия формы «общих данных пациента» для дальнейшего выбора действий оказания медицинской помощи, основными функциями которой являются:

- просмотр электронной медицинской карты, истории болезни, показателей здоровья и другой медицинской информации;
- просмотр контактных данных, адреса проживания, место работы;
- создание медицинских документов.

← → ☆ Список зарегистрированных рисков ухудшения состояния (CC3) 🔗 ⓘ ✕

Ознакомился Открыть ОДП Еще ▾

Документ результата	Пациент	Диагноз	Оценка	Вид	Значение
Регистрация результатов прогнозирования ухудшения...	Т...	I25.1	Есть риск	Общий белок	72,1
				LDH (Лактатдегидрогеназа)	
				PDW-SD (Шир. распр. PLT по V)	16
				Эритроциты(RBC)	4,58
				Sp	140
				ВАС# (Базофилы,#)	0,03
				NT-proBNP (N-проМНП, экспресс)	
				Гемоглобин(HGB)	136
				NEU# (Нейтрофилы,#)	5,93
				Группа крови	
				TnI (ТропонинI, экспресс-тест)	24
				Рост	176
				DBIL (Билирубин прямой)	1,7
				Эозинофилы	
				С-реактивный белок (качественный)	1
				Прозрачность	0 Отрицате
				Средний объем эритроцита(MCV)	87,2
				Нитриты	1 012
				Лейкоциты(WBC)	8,96
				Трикуспидальная регургитация	
				Ядерные эритроциты (NRBC)	
				Базофилы, % (микроскопия)	0,3
				Эозинофилы, % (микроскопия)	1
				Dia	90
				LYM# (Лимфоциты,#)	2,3
				K (Калий)	4,7
				Уробилиноген	0,0 единиц
				Белок в моче (колич.)	0,00 0-1
				LDL (ЛПНП)	
				Резус-фенотипирование	
				MPV (Сред. объем тромбоцита)	9,4
				Лимфоциты, % (микроскопия)	25,7
				PTT (Протромбиновое время)	12,1
				Аортальный стеноз	
				Фибриноген	3,88
				Вес	112
				Моноциты, % (микроскопия)	6,8
				PLT# (Тромбоциты,#)	213
				Сред. сод. гемоглобина эритроците (MCH)	29,8
				INR (МНО)	1,14
				Кислотно-щелочная реакция	
				Кетоны	Прозрачная

Рисунок 44 - Форма списка "текущих дел" выявленных рисков ухудшения состояния здоровья

4.3 Вывод

Спроектирована и разработана подсистема прогнозирования ухудшения состояния здоровья пациентов с СЗЗ для МИС «1С:Медицина.Регион» в виде расширения. Описаны и созданы основные объекты конфигурации, реализован программный код для выполнения основных возможностей подсистемы, часто используемый код вынесен в объекты конфигурации «Общий модуль».

Переопределена работа процессов регистрации результатов лабораторных исследований и формирования «текущих дел» пользователя. Настроено движение данных между документами и регистрами сведений.

Проверено взаимодействие между разработанной подсистемой и веб-сервисом.

5 ЗАКЛЮЧЕНИЕ

В рамках выпускной квалификационной работы был проведен анализ видов показателей здоровья, которые используются лечащими врачами для определения необходимой медицинской помощи пациентам, больным ССЗ.

Была создана и обучена модель бинарной классификации для проведения прогнозирования наличия риска на основании переданных значений показателей здоровья.

Для взаимодействия модели со сторонними системами был разработан веб-сервис, предоставляющий возможности получения списка видов показателей здоровья для прогноза и получение оценки риска на основании переданного набора данных.

Со стороны МИС «1С:Медицина.Регион» была разработана подсистема в виде расширения, взаимодействующая с веб-сервисом, предоставляющая возможности оперативного получения оценки риска ухудшения состояния здоровья пациента сразу после регистрации результатов лабораторных исследований. После получения оценок, на начальной странице МИС пользователю отображаются нерассмотренные результаты прогнозирования с оценкой «Есть риск» в виде «текущих дел».

СПИСОК ЛИТЕРАТУРЫ

- [1] Д. Т. М. Ш. Н. С. А.Ф. Файзрахманов, «МАШИННОЕ ОБУЧЕНИЕ В МЕДИЦИНЕ: ЭВОЛЮЦИЯ И ПЕРСПЕКТИВЫ,» *ЮЖНО-СИБИРСКИЙ НАУЧНЫЙ ВЕСТНИК*, р. 7, 2011.
- [2] Ю. В. Нагибина и Л. А. Захарова, «Медико-социальные особенности больных ишемической болезнью сердца и качество жизни,» 2017. [В Интернете]. Available: <https://elibrary.ru/item.asp?id=28860139>.
- [3] Л. Г. Гелис, «Острый коронарный синдром,» [В Интернете]. Available: http://www.cardio.by/statyagelis_page.
- [4] «Региональная программа "Борьба с сердечно-сосудистыми заболеваниями",» 26 Май 2022. [В Интернете]. Available: https://admtyumen.ru/files/upload/OIV/D_zdr/%D0%9F%D1%80%D0%B8%D0%BB%D0%BE%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5_%D0%9F%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B0%20%D0%91%D0%BE%D1%80%D1%8C%D0%B1%D0%B0%20%D1%81%20%D1%81%D0%B5%D1%80%D0%B4%D0%B5%D1%87%D0%BD%D0.
- [5] Ю. Х. Е., «Система компоновки данных,» [В Интернете]. Available: <https://v8.1c.ru/platforma/sistema-komponovki-dannykh/>.
- [6] А. Викулина, «Регламентные и фоновые задания 1С 8.3,» 26 Март 2019. [В Интернете]. Available: <https://wiseadvice-it.ru/o-kompanii/blog/articles/reglamentnye-i-fonovye-zadaniya-1s-8-3/>.
- [7] «Задачи Data Mining. Классификация и кластеризация,» [В Интернете]. Available: <https://intuit.ru/studies/courses/6/6/lecture/166>.
- [8] «Логистическая регрессия,» [В Интернете]. Available: <https://wiki.loginom.ru/articles/logistic-regression.html>.

- [9] «Классификация данных методом k-ближайших соседей,» [В Интернете]. Available: <https://loginom.ru/blog/knn>.
- [10] «Машинное обучение для начинающих: алгоритм случайного леса,» [В Интернете]. Available: <https://proglib.io/p/mashinnoe-obuchenie-dlya-nachinayushchih-algoritm-sluchaynogo-lesa-random-forest-2021-08-12>.
- [11] «Метод опорных векторов (SVM),» [В Интернете]. Available: [https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2_\(SVM\)](https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2_(SVM)).
- [12] «Поиск аномалий,» 19 04 2017. [В Интернете]. Available: <https://alexanderdyakonov.wordpress.com/2017/04/19/%D0%BF%D0%BE%D0%B8%D1%81%D0%BA-%D0%B0%D0%BD%D0%BE%D0%BC%D0%B0%D0%BB%D0%B8%D0%B9-anomaly-detection/>.
- [13] «Решаем задачи машинного обучения с помощью алгоритма градиентного бустинга,» [В Интернете]. Available: <https://proglib.io/p/reshaem-zadachi-mashinnogo-obucheniya-s-pomoshchyu-algoritma-gradientnogo-bustinga-2021-11-25>.
- [14] «Матрица ошибок (Confusion Matrix),» [В Интернете]. Available: <https://www.helenkapatsa.ru/matritsa-oshibok/>.
- [15] Е. Хрусталева, «Расширения 1С,» [В Интернете]. Available: <https://v8.1c.ru/platforma/rasshireniya/>.

1.1 Листинг запроса получения значений показателей здоровья по каждому случаю госпитализации пациентов

```

ВЫБРАТЬ
    Показатели.ВидПоказателяЗдоровья КАК Показатель,
    Показатели.Соответствие КАК Соответствие
ПОМЕСТИТЬ втВидыПоказателейЗдоровья
ИЗ
    &ПоказателиЗдоровья КАК Показатели
;////////////////////////////////////
ВЫБРАТЬ
    ДвижениеПациентовВСтационаре.Пациент КАК Пациент,
    ДвижениеПациентовВСтационаре.МедицинскаяКарта КАК МедицинскаяКарта,
    ДвижениеПациентовВСтационаре.Период КАК ДатаГоспитализации,
    ДвижениеПациентовВСтационаре.Регистратор КАК ДокументГоспитализации
ПОМЕСТИТЬ ВсеГоспитализации
ИЗ
    РегистрСведений.ДвижениеПациентовВСтационаре КАК ДвижениеПациентовВСтационаре
ГДЕ
    //      "Поступление в стационар" не утверждает, что пациент госпитализировался
    ДвижениеПациентовВСтационаре.ТипДвижения =
ЗНАЧЕНИЕ (Перечисление.ТипыДвиженийПациентов.Госпитализация)
    И ДвижениеПациентовВСтационаре.Период МЕЖДУ &ДатаНачала И &ДатаОкончания
;////////////////////////////////////
ВЫБРАТЬ
    ВсеГоспитализации.Пациент КАК Пациент,
    ВсеГоспитализации.МедицинскаяКарта КАК МедицинскаяКарта,
    МАКСИМУМ(ВсеГоспитализации.ДатаГоспитализации) КАК ДатаГоспитализации
ПОМЕСТИТЬ СлучайГоспитализации
ИЗ
    ВсеГоспитализации КАК ВсеГоспитализации
СГРУППИРОВАТЬ ПО
    ВсеГоспитализации.МедицинскаяКарта,
    ВсеГоспитализации.Пациент
;////////////////////////////////////
ВЫБРАТЬ
    ВсеГоспитализации.Пациент КАК Пациент,
    ВсеГоспитализации.МедицинскаяКарта КАК МедицинскаяКарта,
    ВсеГоспитализации.ДатаГоспитализации КАК ДатаГоспитализации,
    МАКСИМУМ(ВсеГоспитализации.ДокументГоспитализации) КАК ДокументГоспитализации
ПОМЕСТИТЬ ГоспитализированныеПациенты
ИЗ
    СлучайГоспитализации КАК СлучайГоспитализации
ВНУТРЕННЕЕ СОЕДИНЕНИЕ ВсеГоспитализации КАК ВсеГоспитализации
    ПО СлучайГоспитализации.Пациент = ВсеГоспитализации.Пациент

```

```

И СлучайГоспитализации.МедицинскаяКарта =
ВсеГоспитализации.МедицинскаяКарта
И СлучайГоспитализации.ДатаГоспитализации =
ВсеГоспитализации.ДатаГоспитализации
СГРУППИРОВАТЬ ПО
    ВсеГоспитализации.Пациент,
    ВсеГоспитализации.МедицинскаяКарта,
    ВсеГоспитализации.ДатаГоспитализации
;////////////////////////////////////
ВЫБРАТЬ
    ГоспитализированныеПациенты.Пациент КАК Пациент,
    ГоспитализированныеПациенты.МедицинскаяКарта КАК МедицинскаяКарта,
    ГоспитализированныеПациенты.ДатаГоспитализации КАК ДатаГоспитализации,
    ГоспитализированныеПациенты.ДокументГоспитализации КАК ДокументГоспитализации,
    Картотека.ИдентификаторПациентаВРМИС КАК ИдентификаторПациентаВРМИС,
    Госпитализация.Основание КАК ДокументПоступления
ПОМЕСТИТЬ ДанныеГоспитализацийПациентов
ИЗ
    ГоспитализированныеПациенты КАК ГоспитализированныеПациенты
    ВНУТРЕННЕЕ СОЕДИНЕНИЕ Справочник.Картотека КАК Картотека
        ПО ГоспитализированныеПациенты.Пациент = Картотека.Ссылка
        И НЕ Картотека.ПометкаУдаления
    ВНУТРЕННЕЕ СОЕДИНЕНИЕ Документ.Госпитализация КАК Госпитализация
        ПО ГоспитализированныеПациенты.ДокументГоспитализации =
Госпитализация.Ссылка
;////////////////////////////////////
ВЫБРАТЬ РАЗЛИЧНЫЕ
    СведенияОДиагнозахПациентов.Пациент КАК Пациент,
    СведенияОДиагнозахПациентов.МедицинскаяКарта КАК МедицинскаяКарта,
    СведенияОДиагнозахПациентов.Диагноз КАК Диагноз,
    СведенияОДиагнозахПациентов.ТипДиагноза КАК ТипДиагноза
ПОМЕСТИТЬ ДиагнозыПациентовСтационара
ИЗ
    РегистрСведений.СведенияОДиагнозахПациентов.СрезПоследних (,
        (Пациент, МедицинскаяКарта) В (
            ВЫБРАТЬ
                ДанныеГоспитализацийПациентов.Пациент,
                ДанныеГоспитализацийПациентов.МедицинскаяКарта
            ИЗ
                ДанныеГоспитализацийПациентов
        )
    ) КАК СведенияОДиагнозахПациентов
ГДЕ
    СведенияОДиагнозахПациентов.Диагноз В (&Диагнозы)
И СведенияОДиагнозахПациентов.ТипДиагноза В (
    ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Направительный) ,
    ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Первичный) ,

```

ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Основной) ,
 ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Заключительный)
)
 И НЕ СведенияОДиагнозахПациентов.Диагноз =
 ЗНАЧЕНИЕ (Справочник.МКБ10.ПустаяСсылка)
 ;////////////////////////////////////
 ВЫБРАТЬ
 ДанныеГоспитализацийПациентов.Пациент КАК Пациент,
 ДанныеГоспитализацийПациентов.МедицинскаяКарта КАК МедицинскаяКарта,
 ДанныеГоспитализацийПациентов.ДатаГоспитализации КАК ДатаГоспитализации,
 ДанныеГоспитализацийПациентов.ДокументГоспитализации КАК ДокументГоспитализации,
 ДанныеГоспитализацийПациентов.ИдентификаторПациентаВРМИС КАК
 ИдентификаторПациентаВРМИС,
 ДанныеГоспитализацийПациентов.ДокументПоступления КАК ДокументПоступления,
 ЕСТЬNULL (ЗаключительныеДиагнозыПациентовСтационара.Диагноз,
 ЕСТЬNULL (ОсновныеДиагнозыПациентовСтационара.Диагноз,
 ЕСТЬNULL (ПервичныеДиагнозыПациентовСтационара.Диагноз,
 НаправительныеДиагнозыПациентовСтационара.Диагноз))) КАК
 Диагноз
 ПОМЕСТИТЬ ДанныеГоспитализацийПациентовПоДиагнозам
 ИЗ
 ДанныеГоспитализацийПациентов КАК ДанныеГоспитализацийПациентов
 ЛЕВОЕ СОЕДИНЕНИЕ ДиагнозыПациентовСтационара КАК
 НаправительныеДиагнозыПациентовСтационара
 ПО ДанныеГоспитализацийПациентов.Пациент =
 НаправительныеДиагнозыПациентовСтационара.Пациент
 И ДанныеГоспитализацийПациентов.МедицинскаяКарта =
 НаправительныеДиагнозыПациентовСтационара.МедицинскаяКарта
 И НаправительныеДиагнозыПациентовСтационара.ТипДиагноза =
 ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Направительный)
 ЛЕВОЕ СОЕДИНЕНИЕ ДиагнозыПациентовСтационара КАК
 ПервичныеДиагнозыПациентовСтационара
 ПО ДанныеГоспитализацийПациентов.Пациент =
 ПервичныеДиагнозыПациентовСтационара.Пациент
 И ДанныеГоспитализацийПациентов.МедицинскаяКарта =
 ПервичныеДиагнозыПациентовСтационара.МедицинскаяКарта
 И ПервичныеДиагнозыПациентовСтационара.ТипДиагноза =
 ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Первичный)
 ЛЕВОЕ СОЕДИНЕНИЕ ДиагнозыПациентовСтационара КАК
 ОсновныеДиагнозыПациентовСтационара
 ПО ДанныеГоспитализацийПациентов.Пациент =
 ОсновныеДиагнозыПациентовСтационара.Пациент
 И ДанныеГоспитализацийПациентов.МедицинскаяКарта =
 ОсновныеДиагнозыПациентовСтационара.МедицинскаяКарта
 И ОсновныеДиагнозыПациентовСтационара.ТипДиагноза =
 ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Основной)

```

ЛЕВОЕ СОЕДИНЕНИЕ ДиагнозыПациентовСтационара КАК
ЗаключительныеДиагнозыПациентовСтационара
      ПО ДанныеГоспитализацийПациентов.Пациент =
ЗаключительныеДиагнозыПациентовСтационара.Пациент
      И ДанныеГоспитализацийПациентов.МедицинскаяКарта =
ЗаключительныеДиагнозыПациентовСтационара.МедицинскаяКарта
      И ЗаключительныеДиагнозыПациентовСтационара.ТипДиагноза =
ЗНАЧЕНИЕ (Перечисление.ТипыДиагнозовПациента.Заключительный)
ГДЕ
      НЕ (НаправительныеДиагнозыПациентовСтационара.Диагноз ЕСТЬ NULL
      И ПервичныеДиагнозыПациентовСтационара.Диагноз ЕСТЬ NULL
      И ОсновныеДиагнозыПациентовСтационара.Диагноз ЕСТЬ NULL
      И ЗаключительныеДиагнозыПациентовСтационара.Диагноз ЕСТЬ NULL)
;////////////////////////////////////
ВЫБРАТЬ
      Картотека.Ссылка КАК Пациент,
      Картотека.ИдентификаторПациентаВРМИС КАК ИдентификаторПациентаВРМИС,
      ДанныеГоспитализацийПациентовПоДиагнозам.МедицинскаяКарта КАК МедицинскаяКарта,
      ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации КАК
ДатаГоспитализации,
      ВЫБОР
          КОГДА ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации =
ДАТАВРЕМЯ(1,1,1) ТОГДА
              ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации
          ИНАЧЕ

      ДОБАВИТЬКДАТЕ (ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации,
МЕСЯЦ, -2)
      КОНЕЦ КАК ДатаНачалаУчетаПоказателей,
      ВЫБОР
          КОГДА ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации =
ДАТАВРЕМЯ(1,1,1) ТОГДА
              ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации
          ИНАЧЕ

      ДОБАВИТЬКДАТЕ (ДанныеГоспитализацийПациентовПоДиагнозам.ДатаГоспитализации, ДЕНЬ,
2)
      КОНЕЦ КАК ДатаОкончанияУчетаПоказателей,
      ДанныеГоспитализацийПациентовПоДиагнозам.ДокументГоспитализации КАК
ДокументГоспитализации,
      ДанныеГоспитализацийПациентовПоДиагнозам.ДокументПоступления КАК
ДокументПоступления
      ПОМЕСТИТЬ ДанныеПациентовВСтационаре
      ИЗ
      ДанныеГоспитализацийПациентовПоДиагнозам КАК
ДанныеГоспитализацийПациентовПоДиагнозам
      ВНУТРЕННЕЕ СОЕДИНЕНИЕ Справочник.Картотека КАК Картотека

```

```

ПО ДанныеГоспитализацийПациентовПоДиагнозам.Пациент = Картотека.Ссылка
И НЕ Картотека.ПометкаУдаления
;////////////////////////////////////
УНИЧТОЖИТЬ ДанныеГоспитализацийПациентовПоДиагнозам
;////////////////////////////////////
ВЫБРАТЬ РАЗЛИЧНЫЕ
    ДанныеПациентовВСтационаре.Пациент КАК Пациент,
    ДанныеПациентовВСтационаре.ИдентификаторПациентаВРМИС КАК
ИдентификаторПациентаВРМИС,
    ДанныеПациентовВСтационаре.МедицинскаяКарта КАК МедицинскаяКарта,
    ДанныеПациентовВСтационаре.ДокументГоспитализации КАК ДокументГоспитализации,
    ДанныеПациентовВСтационаре.ДатаГоспитализации КАК ДатаГоспитализации,
    ДанныеПациентовВСтационаре.ДатаНачалаУчетаПоказателей КАК
ДатаНачалаУчетаПоказателей,
    ДанныеПациентовВСтационаре.ДатаОкончанияУчетаПоказателей КАК
ДатаОкончанияУчетаПоказателей
ПОМЕСТИТЬ ПациентыСтационара
ИЗ
    ДанныеПациентовВСтационаре КАК ДанныеПациентовВСтационаре
;////////////////////////////////////
ВЫБРАТЬ
    ПоказателиЗдоровья.ВидРезультата КАК Показатель,
    ПоказателиЗдоровья.ЗначениеПоказателяЗдоровья КАК ЗначениеПоказателя,
    ПоказателиЗдоровья.Период КАК ДатаЗначения,
    ПоказателиЗдоровья.Пациент КАК Пациент
ПОМЕСТИТЬ вВсеПоказателиЗдоровьяПациентов
ИЗ
    РегистрСведений.ПоказателиЗдоровья КАК ПоказателиЗдоровья
ГДЕ
    ПоказателиЗдоровья.Период МЕЖДУ &ДатаНачалаОтбораПоказателей И
&ДатаОкончанияОтбораПоказателей
    И ПоказателиЗдоровья.Пациент В
        (ВЫБРАТЬ ПациентыСтационара.Пациент ИЗ ПациентыСтационара)
    И ПоказателиЗдоровья.ВидРезультата В
        (ВЫБРАТЬ вВидыПоказателейЗдоровья.Показатель ИЗ
вВидыПоказателейЗдоровья)
    И (ПоказателиЗдоровья.ЗначениеПоказателяЗдоровья <> НЕОПРЕДЕЛЕНО
        И ПоказателиЗдоровья.ЗначениеПоказателяЗдоровья <> ""
        И ПоказателиЗдоровья.ЗначениеПоказателяЗдоровья <>
ЗНАЧЕНИЕ (Справочник.ЗначенияПоказателейЗдоровья.ПустаяСсылка)
)

ОБЪЕДИНИТЬ ВСЕ

ВЫБРАТЬ
    ПоказателиЗдоровьяПриГоспитализации.Вид,
    ПоказателиЗдоровьяПриГоспитализации.Значение КАК ЗначениеПоказателя,

```

```

ДанныеГоспитализацийПациентов.ДатаГоспитализации,
ДанныеГоспитализацийПациентов.Пациент
ИЗ
ДанныеГоспитализацийПациентов КАК ДанныеГоспитализацийПациентов
ВНУТРЕННЕЕ СОЕДИНЕНИЕ Документ.Госпитализация.ПоказателиЗдоровья КАК
ПоказателиЗдоровьяПриГоспитализации
ПО ДанныеГоспитализацийПациентов.ДокументГоспитализации =
ПоказателиЗдоровьяПриГоспитализации.Ссылка
ГДЕ
ПоказателиЗдоровьяПриГоспитализации.Вид В
(ВЫБРАТЬ втВидыПоказателейЗдоровья.Показатель ИЗ
втВидыПоказателейЗдоровья)
И (ПоказателиЗдоровьяПриГоспитализации.Значение <> НЕОПРЕДЕЛЕНО
И ПоказателиЗдоровьяПриГоспитализации.Значение <> """"
И ПоказателиЗдоровьяПриГоспитализации.Значение <>
ЗНАЧЕНИЕ (Справочник.ЗначенияПоказателейЗдоровья.ПустаяСсылка)
)
;////////////////////////////////////
ВЫБРАТЬ
ПациентыСтационара.Пациент КАК Пациент,
ПациентыСтационара.МедицинскаяКарта КАК МедицинскаяКарта,
ПациентыСтационара.ДокументГоспитализации КАК ДокументГоспитализации,
ПациентыСтационара.ДатаГоспитализации КАК ДатаГоспитализации,
ПациентыСтационара.ДатаНачалаУчетаПоказателей КАК ДатаНачалаУчетаПоказателей,
ПациентыСтационара.ДатаОкончанияУчетаПоказателей КАК
ДатаОкончанияУчетаПоказателей,
втВидыПоказателейЗдоровья.Показатель КАК Показатель,
втВидыПоказателейЗдоровья.Соответствие КАК Соответствие
ПОМЕСТИТЬ ВидыПоказателейЗдоровьяГоспитализированныхПациентов
ИЗ
ПациентыСтационара КАК ПациентыСтационара
ЛЕВОЕ СОЕДИНЕНИЕ втВидыПоказателейЗдоровья КАК втВидыПоказателейЗдоровья
ПО ИСТИНА
;////////////////////////////////////
ВЫБРАТЬ
ВидыПоказателейЗдоровьяГоспитализированныхПациентов.МедицинскаяКарта КАК
МедицинскаяКарта,
ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Пациент КАК Пациент,
ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Показатель КАК Показатель,
ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Соответствие КАК
Соответствие,
МАКСИМУМ (втВсеПоказателиЗдоровьяПациентов.ДатаЗначения) КАК ДатаЗначения
ПОМЕСТИТЬ МаксДатыПоказателиЗдоровья
ИЗ
ВидыПоказателейЗдоровьяГоспитализированныхПациентов КАК
ВидыПоказателейЗдоровьяГоспитализированныхПациентов

```

```

ЛЕВОЕ СОЕДИНЕНИЕ втВсеПоказателиЗдоровьяПациентов КАК
втВсеПоказателиЗдоровьяПациентов
    ПО ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Пациент =
втВсеПоказателиЗдоровьяПациентов.Пациент
    И ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Показатель =
втВсеПоказателиЗдоровьяПациентов.Показатель
    И втВсеПоказателиЗдоровьяПациентов.ДатаЗначения МЕЖДУ
ВидыПоказателейЗдоровьяГоспитализированныхПациентов.ДатаНачалаУчетаПоказателей И
ВидыПоказателейЗдоровьяГоспитализированныхПациентов.ДатаОкончанияУчетаПоказателей
СГРУППИРОВАТЬ ПО
    ВидыПоказателейЗдоровьяГоспитализированныхПациентов.МедицинскаяКарта,
    ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Пациент,
    ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Показатель,
    ВидыПоказателейЗдоровьяГоспитализированныхПациентов.Соответствие
;////////////////////////////////////
УНИЧТОЖИТЬ ВидыПоказателейЗдоровьяГоспитализированныхПациентов
;////////////////////////////////////
ВЫБРАТЬ
    МаксДатыПоказателиЗдоровья.МедицинскаяКарта КАК МедицинскаяКарта,
    МаксДатыПоказателиЗдоровья.Пациент КАК Пациент,
    МаксДатыПоказателиЗдоровья.Соответствие КАК ИДПоказателя,
    МаксДатыПоказателиЗдоровья.ДатаЗначения КАК ДатаЗначения,
    МАКСИМУМ(МаксДатыПоказателиЗдоровья.Показатель) КАК Показатель,
    МАКСИМУМ(втВсеПоказателиЗдоровьяПациентов.ЗначениеПоказателя) КАК
ЗначениеПоказателя
    ИЗ
    МаксДатыПоказателиЗдоровья КАК МаксДатыПоказателиЗдоровья
    ВНУТРЕННЕЕ СОЕДИНЕНИЕ втВсеПоказателиЗдоровьяПациентов КАК
втВсеПоказателиЗдоровьяПациентов
    ПО МаксДатыПоказателиЗдоровья.Пациент =
втВсеПоказателиЗдоровьяПациентов.Пациент
    И МаксДатыПоказателиЗдоровья.Показатель =
втВсеПоказателиЗдоровьяПациентов.Показатель
    И МаксДатыПоказателиЗдоровья.ДатаЗначения =
втВсеПоказателиЗдоровьяПациентов.ДатаЗначения
СГРУППИРОВАТЬ ПО
    МаксДатыПоказателиЗдоровья.МедицинскаяКарта,
    МаксДатыПоказателиЗдоровья.Пациент,
    МаксДатыПоказателиЗдоровья.Соответствие,
    МаксДатыПоказателиЗдоровья.ДатаЗначения

```