За последние несколько десятилетий смертность пациентов с острым коронарным синдромом (ОКС) возросла [1] и стала ведущей причиной смертности во всем мире [2]. По данным Всемирной организации здравоохранения, острый коронарный синдром является основной причиной смерти во всем мире. Ранняя диагностика острого коронарного синдрома и е го прогнозирование очень важны для пациентов с заболеваниями сердца. С другой стороны, очень трудно точно предсказать тождественность острого коронарного синдрома по медицинскому набору данных, поскольку он зависит от множества факторов риска.

В 1960-х годах во Фрамингемском исследовании сердца [4] была выдвинута идея острого коронарного синдрома, и модель прогнозирования острого коронарного синдрома была разделена на два метода, а именно методы, основанные на регрессии, и методы, основанные на машинном обучении. Существует множество моделей прогнозирования риска, основанных на регрессиях, но наиболее распространенными моделями прогнозирования риска для раннего прогнозирования и диагностики основных неблагоприятных сердечно-сосудистых событий являются тромболизис при инфаркте миокарда (TIMI) [5] и Глобальный регистр острых коронарных событий (GRACE) [6], которые используются для прогнозирования оценки риска острого коронарного синдрома. Обе модели используют предыдущую медицинскую карту для изучения и прогнозирования тяжести состояния пациентов, но есть и некоторые недостатки этих старых моделей прогнозирования оценки риска, поскольку они были разработаны и внедрены около 10 лет назад. Эти модели используют несколько человек для прогнозирования риска и предсказывают уровень смертности на основе этих предикторов риска. Существует также больше предикторов, которые могут быть использованы для прогнозирования существования серьезных неблагоприятных сердечно-сосудистых событий (MACE), таких как предыдущая медицинская карта и текущее состояние здоровья пациента.

Существует два метода диагностики и прогнозирования случаев острого коронарного синдрома: клинические методы и модель прогнозирования риска для постановки диагноза. Клиническими методами диагностики острого коронарного синдрома являются ангиография, электрокардиограмма (ЭКГ), холтеровское мониторирование, эхокардиограмма, нагрузочный тест, катетеризация сердца, компьютерная томография сердца (КТ) и магнитно-резонансная томография сердца (МРТ) [7]. Другой метод заключается в проектировании и разработке моделей прогнозирования рисков для ранней диагностики и прогнозирования ОКС с использованием алгоритмов статистического анализа и машинного обучения.

Алгоритмы машинного обучения повышают точность прогнозирования сердечно-сосудистых заболеваний и предотвращают ненужное лечение [8]. Методы машинного обучения преодолели проблемы традиционных методов, основанных на регрессии, и популярны для диагностики и прогнозирования возникновения MACE. Кроме того, он устраняет типичные проблемы с данными и устраняет отсутствующие значения и выбросы с помощью методов интеллектуального анализа данных. Методы машинного обучения основаны на нелинейных связях и взаимодействиях между несколькими переменными и имеют дело с различными предикторами риска для точного прогнозирования риска пациентов. В этом исследовании также изучается эффективность методов прогнозирования риска на основе машинного обучения для прогнозирования степени тяжести пациентов с острым коронарным синдромом. Johnson et al. [9] отметили важность алгоритмов машинного обучения для прогнозирования и диагностики сердечно-сосудистых заболеваний. Тем не менее, методы, основанные на машинном обучении, имеют ряд сложных проблем для прогнозирования случаев MACE в группах ИМпST и NSTEMI у пациентов с острым коронарным синдромом, а именно: Во-первых, не существует специфического машинного обучения или ансамблевого подхода, что дает хорошие результаты для прогнозирования и работы с такого рода клиническими наборами данных. Кроме того, необходимо определить указанные предикторы, влияющие на возникновение острого коронарного синдрома и оказывающие большое влияние на MACE. К сожалению, старые модели прогнозирования в основном основаны на регрессии или их точность колеблется от 65 до 84% [10]. Кроме того, эти модели зависят от нескольких факторов риска. Существуют и другие факторы риска, которые оказывают большее влияние на возникновение острых коронарных синдромов. Кроме того, есть и другие факторы, которые мы должны вывести из других атрибутов и которые оказывают большое влияние на острый коронарный синдром.

Таким образом, в данной работе предлагается ансамблевый классификатор на основе машинного обучения с мягким голосованием, который может заниматься ранней диагностикой и прогнозом у пациентов с острым коронарным синдромом и обеспечить наилучший метод борьбы с возникновением сердечных событий. Основной целью данной работы является разработка модели прогнозирования риска для раннего выявления случаев MACE в течение двухлетнего наблюдения после выписки из стационара у пациентов с острым коронарным синдромом. Содержание нашего исследования также можно резюмировать следующим образом.

Во-первых, для экспериментов мы используем набор данных Корейского регистра острого инфаркта миокарда (KAMIR-NIH) [11], который разделен на две подгруппы: ИМпST и ИМпST.

Во-вторых, мы предлагаем классификатор ансамбля мягкого голосования с использованием алгоритмов машинного обучения, таких как случайный лес (RF), дополнительное дерево (ET) и машина градиентного бустинга (GBM), для повышения точности диагностики и прогнозирования случаев MACE [12], таких как сердечная смерть, несердечная смерть, инфаркт миокарда (ИМ), повторное чрескожное коронарное вмешательство (re-PCI) и аортокоронарное шунтирование (АКШ).

В-третьих, мы уточним предикторы риска MACE для групп ИМпST и ИМпST между предыдущими моделями и нашей новой моделью и сравним результаты этих моделей. Наконец, мы сравниваем результаты прогнозирования случаев MACE в группах ИМпST и NSTEMI в течение двухлетнего наблюдения после выписки из больницы между применяемыми методами машинного обучения (RF, ET и GBM) и нашим классификатором ансамбля мягкого голосования **с помощью показателей эффективности: точность, точность, полнота, оценка F1 и площадь под ROC-кривой (AUC).**

В-четверттых, мы проводим отбор наилучших моделей по критерию **АААА**