

РАЗРАБОТКА СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ ДЛЯ ОЦЕНКИ СИСТЕМЫ РЕГУЛЯЦИИ АГРЕГАТНОГО СОСТОЯНИЯ КРОВИ

Е.А. Чурсина

Научный руководитель: Кудинов А.В.
Томский политехнический университет
a.lena.chur@gmail.com

Введение

Нарушения в работе системы регуляции агрегатного состояния крови проявляются в повышении или понижении свертываемости крови и могут стать причиной тромбозов, кровоизлияний, ишемий, инфарктов органов и других опасных для жизни и здоровья патологий.

Одним из методов оценки и мониторинга системы регуляции агрегатного состояния крови является применение технологии низкочастотной пьезотромбоэластографии. Данный метод основывается на анализе изменений свойств крови, происходящих в процессе гемокоагуляции – перехода крови от жидкого состояния в твердо-эластичное [1]. Он лежит в основе работы тромбоэластографа – прибора, который представляет процессы свертывания крови в графическом виде.

Пользователь не всегда может определить диагноз по результатам работы прибора, поэтому данный процесс требует автоматизации.

Целью работы является разработка системы поддержки принятия решения для постановки диагноза после проведения анализа крови с помощью тромбоэластографа.

Задача классификации

Поставленная задача является задачей многоклассовой классификации – это подкатегория методов машинного обучения с учителем, суть которой заключается в идентификации категориальных меток классов для новых экземпляров на основе предыдущих наблюдений [2].

Для решения задачи классификации требуется сформировать тренировочную выборку, выбрать алгоритм машинного обучения. Затем на маркированных тренировочных данных требуется извлечь модель, которая позволяет делать прогнозы о ранее не встречавшихся или будущих данных [3].

Описание метода низкочастотной пьезотромбоэластографии

Во время проведения анализа динамика исследуемого процесса определяется изменениями агрегатного состояния исследуемого образца и регистрируется в виде интегрированной кривой, каждая точка которой A_i определяется состоянием системы в определенный момент времени T_i [1]. Пример интегрированной кривой для реального диагноза представлен на Рисунке 1.

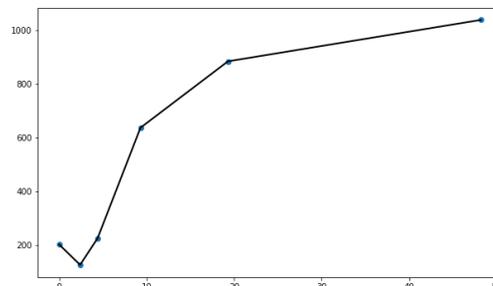


Рис. 1. Пример интегрированной кривой для анализа крови

Кроме этого регистрируются дополнительные параметры, такие как интенсивность контактной коагуляции (ИКК), интенсивность тотального свертывания (ИТС) и другие.

На основе собранных во время проведения анализа данных ставится один из диагнозов: Гиперкоагуляция, Хронометрическая Гиперкоагуляция, Структурная Гиперкоагуляция, Норма в Гиперкоагуляции, Норма Условная, Норма в Гипокоагуляции, Структурная Гипокоагуляция, Хронометрическая Гипокоагуляция, Гипокоагуляция.

Тренировочные данные

В качестве предыдущих наблюдений был взят набор из 1147 записей о реальных анализах крови и их результатах, которые были определены медицинским экспертом. Каждый анализ состоит из 17 параметров и метки соответствующего им диагноза.

Дифференциальный классификатор

Для системы был разработан дифференциальный классификатор, представляющий из себя цепочку использования базовых классификаторов.

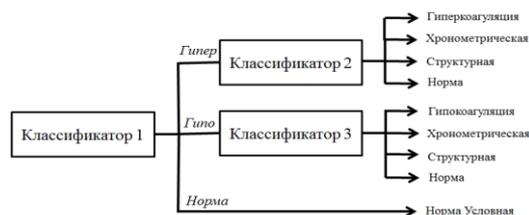


Рис. 2. Схема работы дифференциального классификатора

Классификатор 1 определяет к какому из трех классов (Норма, Гипокоагуляция, Гиперкоагуляция) принадлежит результат анализа. Классификатор 2 работает в случае, если исследование относится к Гиперкоагуляции, и определяет ее вид. Классификатор 3 работает в случае, если исследование относится к Гипокоагуляции, и определяет ее вид.

Реализация подразумевает, что в качестве Классификатора 1, 2 и 3 может быть использован любой переданный в качестве параметра базовый классификатор с любыми параметрами.

Лучшие результаты показывал дифференциальный классификатор, которому в качестве Классификаторов 1, 2 и 3 были переданы классификаторы на основе случайного леса с подобранными отдельно для каждого случая параметрами и наборами признаков, с которыми работает классификатор.

Результаты работы

Точность прогнозирования диагнозов, полученная с помощью кросс-проверки, составляет **91.28%**.

При обучении дифференциального классификатора на тренировочной выборке, составляющей случайные 70% от исходных данных, и проверке работы на тестовой выборке – оставшиеся 30% от исходных данных – были получены следующие результаты:

- Точность прогнозирования – **92.17%**.
- Матрица несоответствий представлена на Рисунке 3.

Ошибки в прогнозировании объясняются несбалансированностью данных и трудной разделимостью таких диагнозов как “Норма в Гиперкоагуляции” и остальных “Гипер” диагнозов и “Норма в Гипокоагуляции” и остальных “Типо” диагнозов. Анализ с результатами “Норма в Гиперкоагуляции” значительно меньше, чем анализ с другими “Гипер” результатами, что мешает модели научиться различать их. По аналогичной причине модели трудно определять диагноз “Норма в Гиперкоагуляции”.

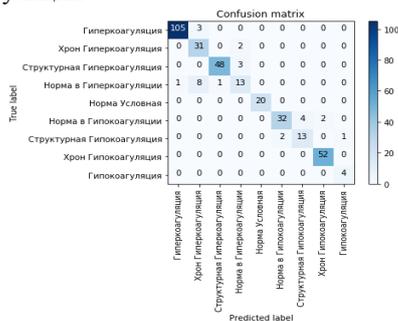


Рис. 3. Матрица несоответствий

Используемые технологии

Для разработки системы поддержки принятия решений был использован стандартный набор инструментов для анализа данных и машинного обучения:

- Язык программирования Python.
- Библиотеки для машинного обучения и работы с данными.

Основные из используемых библиотеки:

- scikit-learn – библиотека, содержащая большое количество алгоритмов машинного обучения [2].
- pandas – библиотека, содержащая инструменты для работы с данными [2].

Заключение

Актуальность оценки и мониторинга системы регуляции агрегатного состояния крови продиктована высокой распространенностью патологических состояний, которые требуют профилактики и лечения.

Разработанная система поддержки принятия решения позволяет автоматизировать процесс постановки диагноза, что делает современную технологию низкочастотной пьезотромбоэластографии более доступной. На данный момент система готовится к внедрению в программное обеспечение тромбоэластографа.

В будущем после получения новых данных качество работы системы может быть улучшено.

Список использованных источников

1. В.В. Удут. Технология низкочастотной пьезотромбоэластографии в оценке гемостатического потенциала // Вестник новых медицинских технологий, электронный журнал - 2016
2. Рашка С. Python и машинное обучение / пер. с англ. А. В. Логунова. М.: ДМК Пресс, 2017. – 418 с.: ил.
3. Davy Cielen, Arno D.B. Meysman, Mohamed Ali Introducing Data Science – Manning Publications Co., 2016.