

Построение нечеткого классификатора для интерпретации результатов оценки свертывающей системы крови у беременных женщин тромбозластографом «Меднорд»

М.Б. Бардамова, В.С. Ковалёв, студенты кафедры КИБЭВС

Научный руководитель И.А. Ходашинский, профессор кафедры КИБЭВС, д.т.н.

г. Томск, ТУСУР, 722bmb@gmail.com

Проект ГПО КИБЭВС-1404 – Нечеткие классификаторы обнаружения вторжений

Введение

Наиболее тяжелым осложнением беременности и родов во всем мире являются акушерские кровотечения. Летальность во время беременности и родов от массивной кровопотери и геморрагического шока составляет от 16 до 36% случаев материнской смертности в разных регионах мира [1] и 18-23% в России [2]. Выявление женщин с нарушениями системы свертывания крови должно проводиться на ранних сроках беременности, и в дальнейшем такие пациентки требуют постоянного контроля со стороны акушеров и гематологов.

Одним из способов контроля состояния гемостаза является оценка свертывающей системы крови с помощью тромбозластографии. На сегодняшний день в Томской области наиболее распространенными тромбозластографами являются приборы производства «Меднорд-техника». Современные модели имеют программное обеспечение, которое в графической и числовой форме представляет результат анализа крови. Однако ПО не способно выдавать заключение по измеренным характеристикам, отсутствует система поддержки принятия решений для врача.

Мы предлагаем для интерпретации результатов использовать классификатор, построенный на нечеткой логике и обученный на данных исследования системы гемостаза. Классификатор способен оказать практическую помощь врачам в оценке результатов и, соответственно, ускорить процесс диагностики и лечения.

Методы нечетких множеств и нечеткой логики широко используются в области классификации данных. К достоинствам нечетких классификаторов можно отнести их хорошую интерпретируемость [3] и отсутствие допущений, необходимых для статистической классификации [4].

Целью данной работы является разработка способа оценки свертывающей системы крови у беременных женщин на основе нечеткого классификатора. Для построения классификаторов были решены следующие задачи: произведен отбор признаков из предоставленной базы анализов пациентов, сгенерированы базы нечетких правил для каждого классификатора, оптимизированы параметры antecedентов (ЕСЛИ-частей) правил.

Методика тромбозластографии

Для сбора данных были проведены исследования на анализаторе реологических свойств крови АРП-01 «МЕДНОРД». В процессе анализа определялись следующие реологические показатели (признаки) [5]: A_n – начальный показатель агрегатного состояния крови; A_r – интенсивность спонтанной агрегации тромбоцитов; r – период реакции; k – время образования сгустка; AM – фибрин-тромбоцитарная константа крови; T – время формирования фибрин-тромбоцитарной структуры сгустка; F – суммарный показатель ретракции и спонтанного лизиса сгустка.

Полученные значения показателей сравнивались с таковыми для здоровой группы лиц, и по характеру их изменения производилась оценка состояния системы гемостаза по трем составляющим: структурные характеристики, хронометрические характеристики, общее состояние. Для каждой из трех характеристик выделялись три основных состояния свертывающей системы: нормальное, гиперкоагуляционный сдвиг, который характеризовался укорочением хронометрических показателей (r , k , T) и увеличением амплитудных характеристик, и гипокоагуляционный сдвиг, выявляемый при удлинении хронометрических характеристик и снижении амплитудных параметров. Все состояния требуют определенную

реакцию со стороны врача-клинициста для коррекции сдвигов и предотвращения потенциальных кровотечений или тромбозов.

Нечеткий классификатор

Нечеткий классификатор строится на основе таблицы наблюдений и базы нечетких правил типа «если-то». В каждом правиле содержится утверждение относительно значений входных переменных и указывается значение выходной переменной в виде метки класса.

Выход классификатора определяет степень принадлежности входных данных из таблицы наблюдений их лингвистическим термам.

Критерий качества классификации определяется с помощью величины, обратной среднеквадратической функции ошибки - точностью классификации. Проблема построения нечеткого классификатора сводится к поиску максимума указанной функции. Для оптимизации параметров термов был использован непрерывный алгоритм гравитационного поиска (GSA_R), а для отбора признаков – дискретный алгоритм гравитационного поиска (GSA_D).

Гравитационный алгоритм оптимизации является метаэвристическим алгоритмом, основанным на фундаментальных законах тяготения. Популяция из входных векторов представляет собой систему частиц, между которыми действуют силы притяжения [6]. Для каждой частицы из популяции на итерации рассчитываются следующие физические характеристики: масса, ускорение, скорость. На последнем шаге происходит обновление позиции частицы путем изменения текущих координат на величину, пропорциональную скорости. Расчеты проводятся до истечения итераций, затем на выход подается вектор с наименьшим значением ошибки классификации.

Входные вектора, подающиеся на вход алгоритму гравитационного поиска, различаются в зависимости от цели оптимизации. Для отбора признаков используются бинарные вектора: «1» означает, что признак включен в обучающую выборку, «0» - признак отсутствует. Задача дискретного алгоритма – отобрать такую обучающую подвыборку, которая обеспечивала бы наибольшую точность и скорость классификации.

Оптимизация вектора параметров antecedентов производится непрерывным алгоритмом. Входом для непрерывного алгоритма гравитационного поиска являются вектора, состоящие из параметров термов. Изменяя параметры и передвигая термы, алгоритм добивается увеличения точности.

Результаты эксперимента

В качестве обучающих данных использовались анализы реологических свойств крови у 500 пациенток, состоящих на наблюдении в ОГАУЗ «Родильный дом №1» г. Томска. На вход классификатору подавалась таблица обучения, состоящая из 8 столбцов: 7 столбцов составляли показатели, измеренные прибором «Меднорд», последний столбец содержал класс, являющийся заключением эксперта-врача. Заключение соответствовало одному из трех состояний системы гемостаза – гиперкоагуляции, нормокоагуляции или гипокоагуляции.

Таблица 1 Пример строки обучающих данных

A_n	A_r	r	k	AM	T	F	Заключение
66	-12	3,7	2,6	787	56,8	15,9	Гиперкоагуляция

На основе обучающих данных были построены три классификатора для определения различных характеристик системы свертывания крови – оценки структурных и хронометрических признаков, а также общего её состояния.

Каждый классификатор тестировался на различных комбинациях признаков, полученных с помощью дискретного гравитационного алгоритмов. В таблицах 2-4 приведены результаты работы классификаторов на тестовых выборках на этапе первичной генерации термов (точность после генерации) и после оптимизации непрерывным гравитационным алгоритмом (точность после оптимизации GSA_R).

Структурные признаки отражают реологические свойства образовавшегося сгустка (таблица 2).

Таблица 2 Результаты тестирования классификатора для оценки структурных признаков

Набор признаков	3 пр., GSA_D r, AM, T	2 пр., GSA_D k, AM	1 пр., GrA AM
Точность после генерации	91,43	93,47	90,57
Точность после оптимизации GSA_R	95,92	95,92	97,84

Хронометрические признаки характеризуют процесс образования тромба (таблица 3).

Таблица 3 Результаты тестирования классификатора для оценки хронометрических признаков

Набор признаков	3 пр., GSA_D A_r , k, T	3 пр., GrA r, k, T	2 пр., GrA r, k
Точность после генерации	56,89	60,89	57,5
Точность после оптимизации GSA_R	82,23	93,42	94,22

Конечным результатом служит комплексная оценка системы гемостаза (таблица 4).

Таблица 4 Результаты тестирования классификатора для оценки общего состояния свертывающей системы

Набор признаков	4 пр., GSA_D r, k, AM, T	4 пр., GSA_D A_r , r, k, AM	3 пр., GrA r, k, AM
Точность после генерации	81,32	84,43	84,85
Точность после оптимизации GSA_R	90,87	91,5	93,15

Заключение

Результаты тестирования нечетких классификаторов позволяют сделать следующие выводы:

нечеткие классификаторы, настроенные приведенными алгоритмами, имеют хорошие способности к обучению (высокий процент правильной классификации на обучающих выборках);

для построения нечеткого классификатора необходима лишь часть реологических показателей крови беременных женщин; информативными являются: r – период реакции; k – время образования сгустка; AM – фибрин-тромбоцитарная константа крови;

предложенный подход может применяться для оценки свертывающей системы крови у беременных женщин.

Список литературы

1. Say L., Chou D., Gemmill A., Tunçalp Ö., Moller A.-B., Daniels J., Gülmezoglu A M., Temmerman M., Alkema L. Global causes of maternal death: a WHO systematic analysis // The Lancet Global Health. 2014. V. 2, N. 6. PP. e323-e333.
2. Куликов А.В., Мартиросян С.В., Обоскалова Т.А. Протокол неотложной помощи при кровотечении в акушерстве. Методические рекомендации. – Екатеринбург, 2010. 38с.
3. Горбунов И.В., Ходашинский И.А. Методы построения трехкритериальных парето-оптимальных нечетких классификаторов // Искусственный интеллект и принятие решений. 2015. № 2. С. 75-87.
4. Scherer R. Multiple Fuzzy Classification Systems // Studies in Fuzziness and Soft Computing, v. 288. – Berlin: Springer-Verlag, 2012. 132 p.
5. Тютрин И.И., Шписман М.Н., Стеценко А.И. Методика исследования и интегральной оценки реологических свойств крови / Актуальные проблемы клинических исследований крови. – Томск STT, 1997. С. 8-18.
6. Rashedi E., Nezamabadi-pour H., Saryazdi S. GSA: A Gravitational Search Algorithm // Information Sciences. 2009. V.179. P. 2232-2248.