



А.В. КУЗНЕЦОВА,

ст. научн. сотр., канд. биол. наук, Институт биохимической физики им. Н.М. Эмануэля РАН,
e-mail: azfor@yandex.ru

О.В. СЕНЬКО,

д.ф.-м.н., ведущ. научн. сотр., ФИЦ «Информатика и управление» им. А.А. Дородницына РАН,
e-mail: senkoov@mail.ru

Ю.О. КУЗНЕЦОВА,

IT-консультант, ООО «Азфорус», e-mail: jul1998@mail.ru

ПРЕОДОЛЕНИЕ ПРОБЛЕМЫ «ЧЕРНОГО ЯЩИКА» ПРИ ИСПОЛЬЗОВАНИИ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В МЕДИЦИНЕ

УДК 004.9

Кузнецова А.В., Сенько О.В., Кузнецова Ю.О. *Преодоление проблемы «черного ящика» при использовании методов машинного обучения в медицине* (Институт биохимической физики им. Н.М. Эмануэля РАН, ФИЦ «Информатика и управление» им. А.А. Дородницына РАН, ООО «Азфорус»)

Аннотация. Предложен интерфейс прогноза в машинном обучении, с использованием метода оптимально достоверных разбиений (ОДР) и модифицированного метода статистически взвешенных синдромов (МСВС). Интерфейс позволяет преодолеть проблему «черного ящика»: иллюстрировать процесс прогнозирования с помощью диаграмм рассеяния, ROC-кривой и ранжирования набора информативных показателей с показом расположения исследуемого объекта.

Ключевые слова: машинное обучение, прогнозирование.

UDC 004.9

Kuznetsova A.V., Senko O.V., Kuznetsova Ju.O. *Black box problem overcoming in medical applications of machine learning* (Emanuel Institute of Biochemical Physics of Russian Academy of Sciences, Azforus, Ltd)

Abstract. A new interface for machine learning predicting models is proposed. Approach is based on optimal valid partitioning (OVP) technique and the modified method of statistically weighted syndromes (LSWR). The interface allows you to overcome the problem of "black box" illustrating prediction process by scatter plots, ROC curves and informative indicators ranking.

Keywords: machine learning, forecasting.

ВВЕДЕНИЕ

Активное распространение компьютерных технологий и доступа к Интернету в мире привели к удвоению объема информации за последние 2 года. К 2020 г. ожидается увеличение объема до 40 зеттабайт, что превосходит прежние прогнозы на 14%. Растет объем данных, которые потенциально могут быть использованы для решения разнообразных задач диагностики и прогнозирования в самых разнообразных отраслях и сферах. Вместе с тем возрастает роль облачных вычислений в управлении «большими данными».



ми» (Big Data) [1]. Это также способствует распространению компьютерных методов решения задач диагностики и прогнозирования, методов машинного обучения (МО) [2]. Выбор технологии работы с данными зависит от качества и объема данных, от поставленной задачи, от ограничений по скорости работы и мощности компьютера. Вместе с тем, существенным препятствием для распространения компьютерных технологий является их непрозрачность, то есть непонятность для конкретного пользователя предлагаемых решений. Эта непрозрачность является существенной в таких областях как медицина, экономика и др. Способы преодоления данной проблемы рассматриваются в рамках настоящей работы.

1.1. Проблема ясности результата в машинном обучении

Современные методы МО, как правило, не требуют априорных предположений о характере вероятностных распределений, не связаны с какими-либо конкретными предположениями о форме разделяющих или аппроксимирующих поверхностей, сохраняют устойчивость и эффективность в условиях высокой размерности. Арсенал средств МО также включает разнообразные методы снижения признакового пространства [3, 5].

Среди большого числа всевозможных технологий МО наибольшую популярность в последние годы приобрели искусственные нейронные сети, метод опорных векторов, решающие деревья и леса [3, 6]. Сохраняют популярность стандартные статистические методы: логистическая регрессия и наивный байесовский классификатор [3].

Существенным недостатком многих перечисленных технологий является их непрозрачность, то есть непонятность для конкретного пользователя предлагаемых решений. Такая непрозрачность является существенной не

только в медицине, но и в экономике и других областях.

Среди перечисленных технологий наибольшую ясность и наглядность предлагаемых решений обеспечивают решающие деревья и байесовские сети [6].

Байесовские сети, представляющие прогноз или диагностическое решение в наглядной легкоинтерпретируемой конечным пользователем форме, получают всё большее распространение. Вместе с тем для эффективного обучения байесовских сетей требуются обучающие выборки очень большого объема, что во многих случаях недостижимо из-за высокой стоимости исследований или ограниченности выборки по естественным причинам (например, в медицине и фармакологии).

Точность решений, получаемых с помощью отдельных решающих деревьев, нередко оказывается существенно ниже точности решений, получаемых с помощью альтернативных методов.

1.2. Информативная визуализация – как преодоление проблемы

Достаточно высокую эффективность вместе с прозрачностью получаемых решений дает метод мультимодельных статистически взвешенных синдромов, являющийся модификацией метода «Статистически взвешенных синдромов», предложенного в [7]. Его эффективность была доказана на достаточно большом числе задач [7–11].

Этот метод, основанный на голосовании по системе двумерных областей в многомерном пространстве, обеспечивает возможность просмотра в одномерном и двумерном режиме локализации объекта и его окружения. Это позволяет сделать вывод о принадлежности данного объекта определенному классу и понять логику принятия решения классификации при голосовании по выделенным областям.





2. Структура и демонстрация программы «Прогноз-М»

Понятие «черный ящик» печально знакомо специалистам по методам машинного обучения (МО). Большинство таких методов позволяют видеть данные только на входе и на выходе (результат МО), не давая никакого представления, на основании чего данный результат был получен. Мы предлагаем интерфейс для Модифицированного метода «Статистически взвешенных синдромов» (МСВС) [8], полностью преодолевший вышеобозначенный недостаток. В предлагаемом интерфейсе работа решающего правила при прогнозировании для конкретного объекта полностью прозрачна и наглядна при визуализации на диаграммах рассеяния – на парах показателей, на ROC-кривой и при мониторинге с выводом весов.

2.1. Структура программы «Прогноз-М»

Преодоление проблемы «черного ящика» позволяет для прогнозируемого объекта не только получить набор наиболее информативных показателей, позволяющих отнести его к одному из сравниваемых классов, но и видеть прогностические веса в областях многомерного пространства, в которые попадают значения исследуемого объекта. Наглядно можно оценить близость данного значения к основной массе наблюдений объектов обучающей выборки, разделенных на классы с помощью выделения разным цветом. Более того, можно создать план поэтапного перехода из класса с неблагоприятным прогнозом в «хороший» класс, если выделить группу показателей, на которые можно воздействовать с целью их изменения. Ранжирование таких «изменяемых» показателей по информативности и весам позволяет в первую очередь обратить внимание на показатели, от которых больше всего зависит результат прогноза для конкретного объекта.

Программа «Прогноз-М» включает следующие этапы обработки входящей информации: сначала производится выбор режима прогноза: режим с обучением (с использованием обучающей выборки) и режим без обучения (на готовом решающем правиле, полученном ранее).

1) Ввод данных конкретного объекта, для которого производится классификация или прогноз.

2) Вывод результата отнесения объекта к определенному классу.

3) Диаграммы рассеяния для информативных показателей – в одномерном и двумерном режимах вывода показателей.

4) мониторинг принадлежности объекта к различным выделенным в многомерном пространстве признаков областям:

- 4.1 – режим вывода информативных показателей в порядке их расположения в базе данных;
- 4.2 – режим вывода ранжированных по информативности показателей.

5) ROC-кривая, с отмеченным на ней положением для конкретного объекта.

6) Локализация объекта в выделенных областях с конкретизацией преобладания объектов того или иного класса в них.

7) Вывод ранжированных весов для выделенных областей, в которые попадает исследуемый объект.

8) Вывод по информативности и весам списка изменяемых показателей, в которых значение объекта попадает в область преобладания «плохого» класса, для выработки стратегии перехода в область с преобладанием «хорошего» класса.

2.2. Демонстрация программы «Прогноз-М» на медицинском примере

Резкий рост интереса к использованию методов МО в медицине связан с повсеместным распространением электронных источников болезни, распространением технологий



распределённых и облачных вычислений. Технологии МО позволяют строить оптимальные диагностические алгоритмы, опираясь на эмпирические закономерности, связывающие диагноз или исход с наблюдаемыми клиническими, лабораторными или инструментальными показателями, которые принято называть признаками. Построение оптимальных алгоритмов производится в автоматическом или полуавтоматическом режиме.

В качестве примера применения методов МО в медицине рассмотрим прототип интерфейса для прогностической модели, позволяющей оценить риск повторного случая возникновения коронарного синдрома в первые полгода после перенесенного обострения ишемической болезни сердца (ИБС). Из большого числа известных методов (логистическая регрессия, деревья решений, нейронные сети и др.) отобран подход, давший наибольшую точность прогноза: метод оптимально достоверных разбиений (ОДР) [9, 10]

и модифицированный метод статистически взвешенных синдромов (МСВС). $AUC = 0,72$.

В программу «Прогноз-М» из 380-ти показателей из базы данных вошел только 41 наиболее информативный показатель. Отбор показателей осуществляли таким образом, чтобы результат на скользящем контроле был наилучшим. Для того, чтобы сделать прогноз для конкретного пациента в таблицу последовательно вводят значения 41 показателя (рис. 1). Для ускорения работы программы предполагается использование режима быстрого прогноза – без обучения, поскольку ранее полученное при машинном обучении решающее правило работает самостоятельно. Режим с обучением также предусмотрен в программе и может быть задействован, но он требует более длительного времени и включается при пополнении обучающей выборки, в результате чего решающее правило может измениться. В этом случае вывод ROC-кривой также должен быть скорректирован.

ПРОГНОЗ-М

Прогнозирование | Инструменты

№/№	Название показателя	Значение	Интервал допустимых значений
3	им. безропуха	2	1-2
4	Стенокардия	2	1-2
5	ФК_стенокардия	3	1-3
6	Дислипиды_ИС	2	1-2
7	Терминальный инфаркт	2	1-2
8	Физическая активность	3	1-3
9	Принят Длительный АТТ	1	1-2
10	Принят Длительный АТТ	1	1-2
11	Принят препарат на 10	2	1-2
12	Принят Длительный на 10	2	1-2
13	креатинин	156	47-159
14	КДРТЖК_фактор	66	32-72
15	КСРЖК_фактор	45	20-57
16	КСОЖК_фактор	217	9-247
17	ПикА	1,19	0,1-1,59
18	СЮВ	116	16-143
19	коэффициент GR	2	0-3
20	инвалидность GR	2	0-3
21	TNF_308_2,0	2	1-2
22	TNF_308_3,0	1	1-2
23	БокВерхСегм_3,0	2	1-2
24	БокСегментЛевСегм_2,0	1	1-2
25	БокСегментВерхСегм_3,0	2	1-2
26	БокСегментСредСегм_2,0	1	1-2
27	БокСегментСредСегм_4,0	2	1-2
28	ЗадБазСегм_3,0	1	1-2
29	ЗадБокБазСегм_2,0	2	1-2
30	ЗадБокБазСегм_2,0	1	1-2
31	ЗадБокБазСегм_3,0	2	1-2
32	ЗадБокБазСегм_2,0	1	1-2
33	ЗадБокБазСегм_2,0	2	1-2
34	МокБокБазСегм_1,0	2	1-2
35	Образование_2,0	1	1-2
36	ПерВертСегм_1,0	2	1-2
37	ПерВертБокБазСегм_1,0	2	1-2
38	ПерВертБокБазСегм_1,0	1	1-2
39	Работа_1,0	2	1-2
40	Работа_3,0	2	1-2
41	CirculatoryFailure	1	0-2

Количество показателей: 41

☒ ГОТОВО!

Результат распознавания:

Пациент относится к первому классу

Первый класс - высокая вероятность риска повторения Коронарного Синдрома в первые полгода после выписки из стационара.

Результат прогнозирования носит вероятностный характер и является рекомендательным. Точность прогноза 72 %. Ответственность за принятие решения несет лечащий врач.

Рис. 1. Вид интерфейса программы «Прогноз» – ввод данных.



Инструменты информативной визуализации представляют собой диаграммы рассеяния (одномерные и двумерные), ROC-кривую и мониторинг, выводящий информативные показатели в порядке их расположения в базе данных или ранжированные показатели по убыванию информативности. Парный вывод показателей на диаграммы рассеяния с обозначенными границами оптимальных разбиений позволяет видеть достоверное преобладание значений 1 и 2 классов в базовых множествах (квадрантах – областях, ограниченных границами разбиений). Обозначения для классов отличаются цветом.

Значения пациента, для которого производится прогноз, выделены на диаграммах более крупным значком, окрашенным в зависимости от класса, к которому пациент отнесен при прогнозировании (рис. 2). На ROC-кривой показано аналогичным символом значение для пациента. Чем ближе оно к верхней правой части кривой, тем с большей вероятностью пациент относится к первому классу (неблаго-

приятный прогноз). Чем ближе положение на кривой к нижнему левому значению, тем прогноз лучше. При изменении значений показателей у пациента в процессе лечения – в динамике – результат прогноза может также изменяться. В этом случае можно будет видеть перемещение положения символа для пациента вдоль ROC-кривой.

Мониторинг позволяет также закрашивать базовые множества в зависимости от достоверного преобладания там объектов 1-го или 2-го классов. Под диаграммой приведено количество объектов по классам в каждом базовом множестве и их процентное соотношение. В случае попадания значения пациента, для которого производится прогноз, в базовое множество с преобладанием значений показателей объектов 1 класса можно видеть, насколько удален символ пациента от границы, разделяющей два класса.

Для базового множества, в которое попало значение для исследуемого объекта,

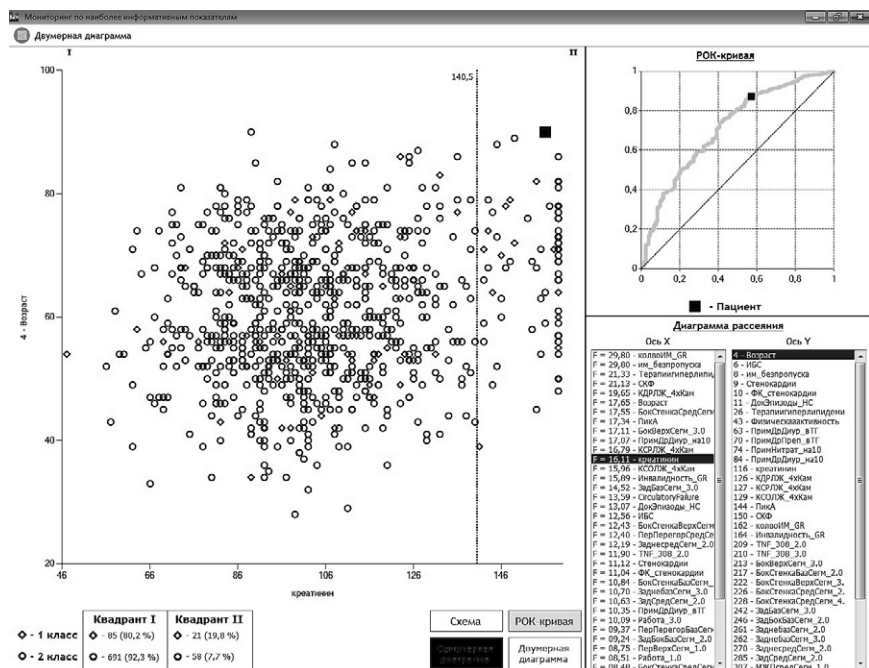


Рис. 2. Мониторинг, одномерные разбиения.
Наиболее информативные показатели.



выводится значение веса для 1 класса. Ранжирование по весам в базовых множествах позволяет также обратить внимание на наиболее важные сочетания показателей, влияющие на результат прогнозирования.

Достижение прозрачности алгоритмов распознавания открывает путь к решению задачи определения объёма и направления коррекции значений прогностических показателей с целью перевода случая, для которого прогноз с использованием алгоритма распознавания неблагоприятен, в группу с благоприятным прогнозом. Можно предположить, что такая коррекция приведёт к существенному увеличению вероятности реализации благоприятного прогноза.

Для решения поставленной задачи может быть предложен подход, включающий выбор тех информативных показателей, значения которых соответствуют областям с высокой долей случаев неблагоприятного исхода. Далее проводится анализ отобранных показателей с оценкой возможности необходимой коррекции. Среди принципиально корректируемых показателей выделяется

подмножество, коррекция которого приводит к значительному повышению оценок вероятности благоприятного прогноза, а стоимость коррекции является по возможности минимальной.

Вывод отчета в файл происходит по нажатию одной кнопки. Список информативных показателей, ранжированных по функционалу (X-квадрат), и весов за 1 класс по базовым множествам сохраняется в документ. Могут быть выведены и все необходимые диаграммы рассеяния.

3. Заключение

Аналогичные программные интерфейсы для прогнозирования могут помочь при машинном обучении во многих сферах: в медико-биологических исследованиях, информатизации здравоохранения, фармакологии и прочих.

Совершенно новым и оригинальным в данном программном продукте является раздел, позволяющий создать пошаговый план перехода объекта из неблагоприятной группы в благоприятную. Для этого в ранжированных по информативности показателей выявляют

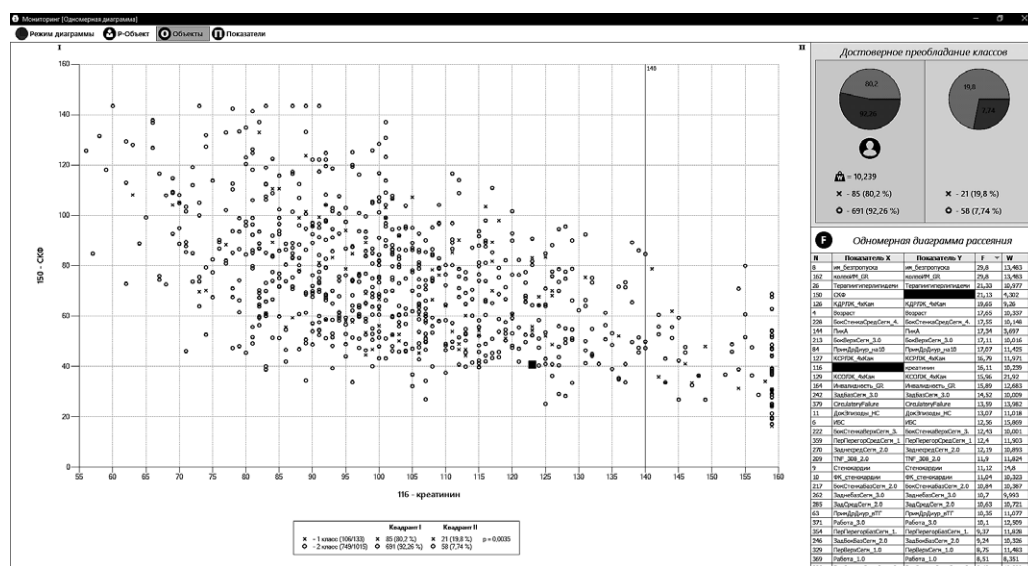


Рис. 3. Вид интерфейса программы «Прогноз» – достоверное преобладание классов.



значения пациента, которые находятся в зоне риска и по которым возможна коррекция, т.е. модифицируемые показатели.

Программа интерфейса «Прогноз-М» полностью подготовлена для работы с самыми разнообразными данными. При постановке новой задачи прогнозирования адаптация алгоритма к новой обучающей выборке займет минимум времени. Удобство использования предлагаемого интерфейса ввиду его наглядности и полной прозрачности в значительной степени улучшит понимание пользователем (аналитиком) закономерностей, выявленных

в процессе машинного обучения. Это преимущество в свою очередь позволит понять наиболее значимые процессы, приводящие к тому или иному состоянию системы, а также пути перехода в нужное положение в многомерном пространстве признаков.

Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ – проект 17-07-01362. Выражаем благодарность за неоценимую помощь при создании программы С.Н. Лобанову (ООО «Азфорус»).

ЛИТЕРАТУРА



1. Рост объема информации – реалии цифровой вселенной. Журнал «Технологии и средства связи». – 2013. – № 1. – С. 24.
2. *Darcy A.M., Louie A.K., Roberts L.W.* Machine Learning and the Profession of Medicine. *Journal of American Medical Association.* – 2016. – №. 6. – V. 315.
3. *Hastie T., Tibshirani R., Friedman J.* Statistical Learning. Data Mining, Inference, and Prediction. Second Edition. Springer Series in Statistics, Springer. – 2008. – 678 p.
4. *Aguilar S., Almeida L.L., Ruffino-Netto A., Kritski A.L., CQ Mello F. and. Werneck L.* Classification and regression tree (CART) model to predict pulmonary tuberculosis in hospitalized patients. *BMC Pulmonary Medicine*, 2012, 12:40.
5. Marafino HYPERLINK "https://www.ncbi.nlm.nih.gov/pubmed/?term=Marafino%20BJ%5BAuthor%5D&cauthor=true&cauthor_uid=25700665", B.J., Boscardin HYPERLINK "https://www.ncbi.nlm.nih.gov/pubmed/?term=Boscardin%20WJ%5BAuthor%5D&cauthor=true&cauthor_uid=25700665", W.J., Dudley, R.A. Efficient and sparse feature selection for biomedical text classification via the elastic net: Application to ICU risk stratification from nursing notes. *J. Biomed.* HYPERLINK "<https://www.ncbi.nlm.nih.gov/pubmed/25700665>" *Inform.*, 2015. Apr; 54: 114–20.
6. *Cruz-Ramirez N., Acosta-Mesa H.G., Carrillo-Calvet H., Nava-Fernández L.A., arrientos-Martínez R.E.* Diagnosis of breast cancer using Bayesian networks: a case study. *Comput. Biol. Med.*, 2007; 37: 1553–64.
7. *Кузнецов В.А., Сенько О.В., Кузнецова А.В. и др.* Распознавание нечетких систем по методу статистически взвешенных синдромов и его применение для иммуногематологической нормы и хронической патологии. *Химическая физика.* – 1996. – Т. 15. – № 1. – С. 81–100.
8. *Senko O.V., Kuznetsova A.V.* A recognition method based on collective decision making using systems of regularities of various types. *Pattern Recognition and Image Analysis.* – 2010. – V. 20. – № 2. – P. 152–162. doi: 10.1134/S1054661810020069.
9. *Kuznetsova, A.V., Kostomarova, I.V., Senko, O.V.* Modification of the method of optimal valid partitioning for comparison of patterns related to the occurrence of ischemic stroke in two groups of patients. *Pattern Recognition and Image Analysis.* – 2014. – V. 24. – № 1. – P. 114–123. doi: 10.1134/S105466181401009X.
10. *Senko O.V., Kuznetsova A.V.* The Optimal Valid Partitioning Procedures. *InterStat.* 2006. April. № 2.
11. *Гулиев Р.Р., Сенько О.В., Затейщиков Д.А. и др.* Применение оптимальных разбиений для многопараметрического анализа данных в клинических исследованиях. *Математическая биология и биоинформатика.* – 2016. – Т. 11. – № 1. – С. 46–63. doi: 10.17537/2016.11.46