

А.А. Болотов¹, В.Г. Барчуков¹, А.С. Галузин¹, А.М. Тен², О.В. Онопоченко²

ИСПОЛЬЗОВАНИЕ ГЕМАТОЛОГИЧЕСКИХ ПОКАЗАТЕЛЕЙ ДЛЯ ПРОГНОЗА РЕЗУЛЬТАТОВ СОРЕВНОВАНИЙ СПОРТСМЕНОВ ПЛОВЦОВ НА ОСНОВЕ ВЕРОЯТНОСТНЫХ ОЦЕНОК

¹ФГБУ ГНЦ ФМБЦ им. А.И. Бурназяна ФМБА России, Москва²ФГБУ ФНКЦСМ ФМБА России, Москва

Контактное лицо: Болотов Александр Александрович: abolotov@bk.ru

Резюме

Цель: В статье рассматриваются вероятностные подходы к анализу гематологических показателей спортсменов в плавании, их использования для оценки функционального состояния с целью прогнозирования успешности выступлений на соревнованиях и создания компьютерной системы для оценки прогноза успешности выступлений на основе наиболее значимых гематологических признаков.

Материал и методы: На базе вероятностных подходов предложена методика оценки влияния гематологических показателей на определение функционального состояния (в целях прогноза успешности выступлений) спортсменов на соревнованиях. Рассматриваются два вида оценок значимости показателей: интегральная оценка и градационная оценка влияния значений каждого признака на результат прогноза. Интегральная оценка значимости предназначена для отбора признаков, наиболее сильно влияющих на результат прогноза, а градационная оценка позволяет определить влияние значений каждого показателя на прогноз успешности и служит для объяснения результата прогноза.

Результаты: Применение оценок интегральной значимости признаков позволило уменьшить число гематологических показателей крови с 18 до 8 для разработки компьютерной системы оценки прогноза успешности выступления спортсменов на соревнованиях. Предложена методика перевода непрерывных показателей в дискретную шкалу с тремя градациями на основе критерия максимума интегральной значимости показателя. Приведена оценка влияния каждой градации признака на вероятность успешного прогноза. Разработана компьютерная система прогноза успешности выступления спортсменов. Предложена методика оценки качества заключений системы прогноза успешности выступления спортсменов на основе методов Монте-Карло.

Область применения: Результаты работы имеют прикладное значение для построения компьютерных систем оценки функционального состояния и прогноза успешности выступлений спортсменов с целью улучшения качества отбора кандидатов как в командных дисциплинах, так и в личных состязаниях спортсменов.

Выводы: Применение оценок интегральной значимости признаков позволило уменьшить число гематологических показателей для разработки системы оценки прогноза успешности выступления спортсменов на соревнованиях. Результаты статистического тестирования методики оценки качества заключений системы прогноза успешности выступления спортсменов на подходах Монте-Карло показали достаточно высокую точность (около 93%) системы прогноза на имеющемся объеме фактических данных, что позволяет сделать вывод о ее работоспособности и возможности дальнейшего развития.

Ключевые слова: вероятностная оценка информативности признаков, гематологические показатели, метод Монте-Карло, прогноз успешности, функциональное состояние спортсменов

Для цитирования: Болотов А.А., Барчуков В.Г., Галузин А.С., Тен А.М., Онопоченко О.В. Использование гематологических показателей для прогноза результатов соревнований спортсменов пловцов на основе вероятностных оценок // Клинический вестник ФМБЦ им. А.И. Бурназяна 2023. №4. С. 10–18. DOI: 10.33266/2782-6430-2023-4-10-18

А.А. Bolotov¹, V.G. Barchukov¹, A.S. Galuzin¹, A.M. Ten², O.V. Onopchenco²

Using Hematological Indicators to Predict the Results of Competitions of Swimmers Based on Probabilistic Estimation

¹International Office, State Research Center - Burnasyan Federal Medical Biophysical Center of Federal Medical Biological Agency, Moscow, Russia²Federal Research and Clinical Center of Sports Medicine and Rehabilitation of Federal Medical Biological Agency, Moscow, Russia

Contact person: Bolotov Aleksandr Aleksandrovich: abolotov@bk.ru

Abstract

Purpose: The article discusses probabilistic approaches to the analysis of hematological indicators of athletes in swimming, their use for assessing the functional state in order to predict the success of performances in competitions and the creation of a computer system for assessing the prediction of successful performances based on the most significant hematological signs.

Methodology of the work: On the basis of probabilistic approaches, a method for assessing the influence of hematological parameters on the determination of the functional state (in order to predict the success of performances) of athletes in competitions is proposed. Two assessments of the significance of indicators are considered: an integral assessment and an assessment of the influence of the values of each feature on the forecast result. The integral significance assessment is designed to select the features that most strongly affect the prediction result, and the assessment of the influence of the values of each feature allows you to determine the impact of the values of each indicator on the success prediction and serves to explain the prediction result.

Results: The use of estimates of the integral significance of signs made it possible to reduce the number of hematological blood parameters from 18 to 8 in order to develop a computer system for evaluating the prediction of the success of athletes

in competitions. A technique for converting continuous indicators into a discrete scale with three gradations based on the criterion of the maximum integral significance of the indicator is proposed. An assessment of the influence of each feature gradation on the probability of a successful forecast is given. A computer system for predicting the success of the performance of athletes has been developed. A technique for assessing the quality of the conclusions of the system for predicting the success of the performance of athletes based on the Monte-Carlo methods is proposed.

Scope of application of the results: The results of the work are of applied importance for the construction of computer systems for assessing the functional state and predicting the success of athletes' performances in order to improve the quality of the selection of candidates both in team disciplines and in individual competitions of athletes.

Conclusions: The use of assessments of the integral significance of signs made it possible to reduce the number of hematological indicators for the development of a system for assessing the prediction of the success of athletes' performance in competitions. The results of statistical testing of the methodology for assessing the quality of the conclusions of the system for predicting the success of the performance of athletes on Monte-Carlo approaches showed a fairly high accuracy (about 93%) of the prediction system on the available volume of actual data, which allows us to conclude that it is working and the possibility of further development.

Keywords: *functional state of athletes, hematological parameters, Monte-Carlo method, probabilistic assessment of the information content of signs, success forecast*

For citation: Bolotov AA, Barchukov VG, Galuzin AS, Ten A.M, Onopchenko OV. Using Hematological Indicators to Predict the Results of Competitions of Swimmers Based on Probabilistic Estimation. A.I. Burnasyan Federal Medical Biophysical Center Clinical Bulletin. 2023.4:10-18. (In Russian) DOI: 10.33266/2782-6430-2023-4-10-18

Введение

В настоящее время развитие спорта характеризуется все более возрастающей напряженностью тренировочной и соревновательной деятельности. Планирование подготовки спортсменов к соревнованиям осуществляется с учётом годового плана-графика тренировочных занятий. Основой для планирования нагрузок в годичном цикле являются сроки проведения соревнований (контрольные, отборочные, основные). Начиная с базового этапа многолетней спортивной подготовки тренировочные нагрузки пловцов, распределяются на два, три или четыре макроцикла. Чем выше спортивная квалификация пловцов, тем в большей степени выражена волнообразность динамики нагрузки. В макроциклах обычно выделяются подготовительный, соревновательный и переходный периоды.

Тренировочный процесс направлен на гармоничное развитие систем организма, обеспечивающих успешное выступление спортсмена на соревнованиях. Это достигается эффективным использованием методов тренировок, разрешенных медикаментозных, физиотерапевтических, психологических и других воздействий, способных вызвать предельную мобилизацию функциональных возможностей организма пловца в соревновательный период. Ранее нами для подготовительного периода в рамках тренировочных сборов на базе системы интеллектуальной поддержки тренера (СИПТ) [1] были предложены, основанные на данных ступенчатого теста пловцов, модели, позволяющие оптимизировать тренировочный процесс на основе анализа биохимических показателей крови спортсменов.

Однако наряду с базовой тренировочной подготовкой актуальным при выступлении на соревнованиях является оценка функционального состояния спортсмена-пловца и его готовность показать максимальный результат. При этом оценка состояния спортсмена должна быть для него быстрой, малозатратной и без отвлечения от сорев-

новательного процесса. Имеющиеся литературные данные свидетельствуют о том, что инструментом, позволяющим оценить готовность пловца к решению этой задачи, являются его гематологические показатели [2, 3]. С учетом этого было решено оценивать функциональное состояние спортсмена на основе гематологических показателей крови, определенных при заборе крови накануне соревнований. Гематологические показатели позволяют одновременно оценить состояние кислородотранспортной (гемоглобин, эритроциты, гематокрит и производные) и иммунной системы (лейкоциты и их подгруппы). Кислородотранспортная система связана с энергообеспечением, что важно в ходе прохождения дистанции с ускорением на финише, а также для восстановления при повторных стартах в течение соревнований. Достаточное количество энергии также обеспечивает эффективную работу иммунной системы, которая как известно страдает на пике тренировочного процесса, что проявляется увеличением простудных заболеваний, приводящих к срыву стартов по причине болезни [2, 3].

Целью данной работы является исследование возможности использования гематологических показателей спортсменов-пловцов в соревновательный период для оценки функционального состояния и формирования прогноза успешности их выступлений для отбора лучших кандидатов в командных и личных заплывах на этих соревнованиях.

Материалы и методы

Для оценки функционального состояния спортсменов с целью прогноза успешности их выступлений были собраны данные по 18 гематологическим показателям пловцов с оценкой успешности/неуспешности выступления спортсменов, полученных в период проведения соревнований.

Гематологические показатели были полученных с помощью биохимического анализатора «Chem-

Well-Т», производства Awareness Technology (США) [4]. Это полностью автоматизированный турбидиметрический биохимический анализатор, управляемый персональным компьютером, позволяющий проводить полный спектр биохимических и иммунотурбидиметрических анализов до 100 тестов в час. Принцип действия анализатора основан на измерении оптической плотности жидкой пробы и последующем пересчете с помощью встроенных программ полученного значения в необходимый параметр (концентрацию) лабораторного теста.

Функциональное состояние спортсменов будем оценивать показателем успешности их выступления на соревнованиях. Оценка успешности проводилась на основе сравнения наилучших результатов спортсменов за предыдущий сезон с лучшими результатами конкретных соревнований по очкам (баллам) калькулятора FINA (ФИНА), позволяющим давать единую оценку результатов соревнований по различным дисциплинам плавания [5]. Если результаты данных соревнований спортсмена были лучше результатов предыдущего сезона, то результат считался успешным, в противном случае – неуспешным. Кровь у спортсменов забиралась с утра до завтрака перед соревнованиями.

В фактические данные вошло 100 случаев выступления пловцов мужчин - спортсменов сборной РФ (возраст от 18 до 27 лет) на соревнованиях с данными гематологических показателей и известным результатом успешности/неуспешности их выступлений, список которых приведен в таблице 1.

Для оценки прогноза успешности выступления спортсменов на соревнованиях необходимо определить относительное влияние гематологических показателей на прогноз успешности/неуспешности выступления спортсменов, выделить наиболее значимые гематологические показатели спортсменов на имеющихся фактических данных, на их основе разработать систему (программное обеспечение) расчета прогноза успешности и определить точность ее работы.

Величину относительного влияния гематологических показателей на прогноз успешности/неуспешности выступления спортсменов будем оценивать с помощью интегральной оценки значимости показателя, получаемой на основе вероятностных подходов, описанных в [6]. Эти подходы предполагают использование непрерывных и качественных показателей в дискретной форме: признак-градация. По имеющемуся набору фактических данных гематологических показателей спортсменов вычислялись оценки условных вероятностей каждой градации признаков (показателей) для двух классов решений (успешный/неуспешный), т.е. частота присутствия данной градации признака для каждого класса решений в фактических данных.

Также вычислялась априорная вероятность классов решений, т.е. частота появления каждого класса решений в фактических данных. Далее по этим величинам вычислялась интегральная оценка значимости всех показателей (признаков) [6].

Рассмотрим такой подход подробнее, введем следующие обозначения:

Признак Si [1..I] – означает число признаков от 1 до I.

Таблица 1

Список гематологических показателей спортсменов
List of hematological indicators of athletes

№ признака	Наименование	Название показателя (единица измерения)	Мин. - Макс. значения показателя в фактических данных	Значимость*
1	WBC	Лейкоциты ($\times 10^9/\text{л}$)	3,02 - 11,53	1,16
2	RBC	Эритроциты ($\times 10^{12}/\text{л}$)	4,32 - 7	1,15
3	HGB	Гемоглобин (г/л)	121 - 175	1,18
4	HCT	Гематокрит (%)	34,81 - 62,95	1,03
5	MCV	Средний объем эритроцита (фл)	74 - 91	1,37
6	MCH	Среднее содержание гемоглобина в эритроците (пг)	21,8 - 33	1,24
7	MCHC	Средняя концентрация гемоглобина в эритроцитной массе (г/л)	240 - 370	1,21
8	PLT	Тромбоциты ($\times 10^9/\text{л}$)	100 - 310	1,15
9	PCT	Тромбоциты (%)	0,02 - 0,22	1,12
10	MPV	Средний объем тромбоцитов (фл)	6,7 - 10,4	1,1
11	PDWc	Индекс распределения тромбоцитов (%)	35,3 - 45,2	1,07
12	RDWc	Распределение тромбоцитов по их объему (%)	11,8 - 15,8	1,25
13	LYM	Лимфоциты ($\times 10^3/\text{мл}$)	1,02 - 3,49	1,17
14	MON	Моноциты ($\times 10^9/\text{л}$)	0,24 - 1,24	1,07
15	GRA	Гранулоциты ($\times 10^9/\text{л}$)	1,27 - 8,16	1,08
16	LYM%	Лимфоциты (%)	19,3 - 57,2	1,21
17	MON%	Моноциты (%)	6,3 - 22,4	1,08
18	GRA%	Гранулоциты (%)	33,1 - 71,8	1,18

Примечание *-Жирным выделены наиболее значимые показатели

Градация признака k $[1..g]$ – означает число градаций признака от 1 до g .

Класс решений D_j $[1..J]$ – означает число (класс) решений от 1 до J .

Тогда условная вероятность i признака с градацией k для класса j :

$$P_{ij}^k(S_i^k/D_j) = N_{ij}^k / N_j$$

где N_{ij}^k – количество случаев с i признаком градацией k , с классом j

N_j – общее количество случаев с классом j

Априорная вероятность классов решений:

$$P(D_j) = N_j / N$$

где N – общее число случаев.

Интегральная информационная значимость признаков определяется следующим образом [6]:

$$Z_i = \sum_{j=1}^J P(D_j) * \sum_{k=1}^g P^2(S_i^k / D_j) / P_{ij}(S_i) \quad (1)$$

где

$$P_{ij}(S_i) = \sum_{k=1}^g P_{ij}^k(S_i^k / D_j) * P(D_j)$$

При этом величина Z_i признака больше 1 и меньше J – числа классов. Z_i зависит от числа используемых градаций признака: как правило, чем больше градаций, тем больше интегральная значимость признака [6]. При этом чем больше Z_i тем больше значимость этого признака и его влияние на результат классификации (прогноза). Для двух классов решений – успешный или неуспешный интегральная оценка значимости Z_i меньше двух. При значениях интегральной оценки значимости признаков близких к 1 такие признаки являются незначимыми и бесполезными для прогноза. Таким образом появляется возможность ранжирования показателей и выбора только тех, которые наиболее сильно влияют на прогноз.

Для оценки влияния значений каждого признака на результат классификации (прогноза) использовалась градационная значимость признака, т.е. информационная значимость градаций показателя Z_{ij}^k , которая показывает во сколько раз увеличивает или уменьшает вероятность успешного прогноза каждая градация признака [6]:

$$Z_{ij}^k = P_{ij}^k(S_i^k/D_j) / P_{ij}(S_i) \quad (2)$$

где

$$P_{ij}(S_i) = \sum_{k=1}^g P_{ij}^k(S_i^k / D_j) * P(D_j)$$

Если $Z_{ij}^k > 1$ – k градация i признака увеличивает вероятность j класса, и наоборот при $Z_{ij}^k < 1$ – уменьшает вероятность j класса

Для перевода непрерывных признаков в дискретные использовался разработанный нами подход на основе критерия максимума интегральной оценки значимости признака. Для его реализации разработан специальный алгоритм, который разбивает весь диапазон непрерывного показателя на несколько десятков интервалов, и начиная с минимального (в цикле) до максимального значения по этим интервалам происходит расчет интегрального показателя значимости. После каждого цикла интервалы переопределяются. Если значение интегрального показателя значимости превышает предыдущее значение цикла, то это значение интегральной оценки значимости и диапазоны (градации) показателя запоминаются. В итоге работы алгоритма получаем оценки максимальной интегральной значимости показателя и оптимальные градации показателя.

Многочисленными психологическими экспериментами показано, что люди ведут себя надежно и непротиворечиво при ограниченном числе возможных путей (классов) решения, числе признаков (показателей) и числе возможных значений для каждого из признаков [7-11]. Поэтому количество используемых признаков в человеко-машинных системах принятия решений должно быть ограничено с позиций человеческой системы переработки информации [8]. Имеющиеся литературные данные [11] свидетельствуют о том, что принятие решений происходит в кратковременной (рабочей) памяти человека. Объем кратковременной памяти ограничен «магическим» числом 7 ± 2 чанка (чанк – фрагмент знаний, хранимый и используемый как единое целое). Ограничение объема кратковременной памяти оказывает существенное влияние на систему переработки информации человеком и именно это ограничение заставляет людей вырабатывать чанки все более емкого и обобщенного характера, т.е. ограниченный объем кратковременной памяти заставляет экспертов приспосабливать классификационные задачи к своим возможностям. Это обстоятельство необходимо учитывать при построении компьютерной системы, обеспечивающей поддержку принятия решений и возможности объяснения выдаваемых решений.

Целесообразным представляется выделение трех градаций из непрерывных показателей, что часто используется в задачах классификации и прогноза, т.к. такой подход хорошо согласуется с форматом суждений специалистов, используемых в анализе и принятии решений, например, 1-ниже нормы, 2-норма, 3-выше нормы.

Для уточнения числа разбиений признака на градации производится 3 группы расчетов:

- С использованием пяти градаций признака;
- С использованием трех градаций;
- С использованием двух градаций признака.

Сравнение данных 1 и 2 группы расчетов позволяет определиться в выборе только тех показателей, которые в обеих группах парно имеют высокие оценки значимости и только их использовать для

описания признакового пространства на базе трех градаций.

Сравнение данных 2 и 3 групп позволяет выделить показатели, которые хорошо аппроксимируются двумя градациями – когда интегральные оценки значимости показателя при 3-х градациях практически совпадают с оценкой значимости при двух градациях, т.е. в этом случае нет необходимости использовать 3 градации. Уменьшение числа градаций признаков позволяет снизить требования к объему обучающей выборке (ОВ) и значительно уменьшает максимально возможный объем полной базы фактов (данных), содержащей все возможные комбинации признаков. Рассмотрим это на примере 2 ситуаций: так при 8 признаках с 3 градациями число комбинаций (сочетаний) различных признаков будет равно $3^8 = 6531$, а 7 признаках с 3 градациями и одним с 2 будет равно $3^7 * 2 = 4374$. Обучающая выборка обычно представляет собой набор данных содержащих от нескольких десятков до нескольких сотен случаев. Так при ОВ=130 случаев она будет представлять 2% от числа возможных комбинаций признаков первой ситуации и 3% от второй. От этого параметра сильно зависит точность алгоритмов распознавания, так по критерию точности аппроксимации разделяющих гиперплоскостей (ТАРГ) их точность растет в виде логарифмической зависимости от объема выборки [12] и при небольших изменениях объема выборки точность алгоритмов сильно изменяется.

Использование небольшого числа признаков с наибольшей значимостью для систем прогноза и классификации, а также перевод непрерывных показателей дискретные (выделение градаций признаков) позволяет приблизить объяснительные возможности компьютерных систем к суждениям и выводам экспертов, упрощает объяснение результатов прогноза т.к. становится более понятен механизм влияния показателя (признака) для лиц, принимающих решения, что важно для человеко-машинных систем принятия решений.

Результаты и обсуждение

С учетом вышеизложенного из 18 показателей были отобраны 8 наиболее значимых (при использовании трех градаций), которые представлены в таблице 1 и выделены жирным шрифтом. Отбор показателей осуществлялся на основе медианного значения оценок интегральной значимости всех признаков и привлечением врачей по спортивной медицине в качестве экспертов с целью анализа информации по интегральной оценке значимости признаков и возможности их использования для прогноза. Показатель «Лимфоциты» был исключен, и вместо него использовался показатель «Лимфоциты %», более значимый для прогноза, поскольку они определяют один и тот же показатель, выраженный в разных единицах измерения. Признаков, которые хорошо аппроксимируются двумя градациями (по критерию максимальной интегральной значимости) не было выявлено. Значимые показате-

тели с их информативностью по трем градациям приведены в таблице 2. По убыванию значений интегральной значимости признаки располагаются в следующем порядке: 4, 6, 7, 8, 3, 1, 5, 2.

Градации гематологического показателя, у которых оценка его значимости больше единицы увеличивают вероятность успешного прогноза и наоборот - градации показателя, у которых оценка значимости меньше единицы, уменьшают вероятность успешного прогноза.

Так при наличии у спортсменов градации №2 будет увеличиваться вероятность успешного прогноза в 1,3 раза по «Лейкоцитам» и в 2,86 раз по «Гранулоциты %», другие значения градаций этих двух признаков будут уменьшать вероятность успешного прогноза. Таким образом, разбиение непрерывных показателей на 3 градации позволяет дать простую и удобную интерпретацию влияния различных показателей на прогноз успешности выступления спортсменов и при этом становится более понятен механизм влияния показателя (признака) для лиц, принимающих решения о возможности участия того или иного спортсмена на данных соревнованиях.

Для наиболее успешного прогноза будет следующая комбинация показателей с их градациями в формате номер показателя с градацией через точку:

1.2; 2.3; 3.2; 4.1; 5.2; 6.2; 7.2; 8.2

А для наиболее неуспешного прогноза:

1.1; 2.1; 3.1; 4.3; 5.3; 6.3; 7.1; 8.1

Для практической реализации и построения системы прогноза использовались подходы и методы на основе технологии МКЛАСС [13, 14]:

- используется дискретная шкала признаков на основе градаций признаков;
- описание случая задается вектором признаков, который состоит из строки длиной равной числу используемых признаков, где градация признака стоит на позиции номера признака в такой строке;
- для прогноза (классификации) используется база знаний (БЗ), которая содержит базу фактов (БФ) и суждения экспертов;
- система ориентирована на построение больших и непротиворечивых БЗ;
- база знаний формируется на основе продукций, задаваемых вектором признаков, т.е. правил «если вектор признаков ... то класс ...»;
- для прогноза используется комитетный (агрегированный) принцип формирования заключений на основе 3 независимых алгоритмов;
- имеется объяснительная компонента получаемых заключений классификации и др.

Обобщенная схема системы прогноза по гематологическим показателям спортсменов представлена на рис.1.

Для компьютерной системы оценки прогноза был сформирован уменьшенный по числу признаков набор фактических данных (НФД) из 100 случаев выступления пловцов мужчин на соревнованиях по 8 признакам и известным результатом успешности/неуспешности их

Таблица 2

Наиболее значимые гематологические показатели по трем градациям с их информативностью
The most significant hematological indicators according to three gradations with their information content

№ признака	Название признака	Градации признака	Информативность градации признака
1	Лейкоциты	1 < 3,587	0,01
		2 3,587 - 6,708	1,3
		3 > 6,708	0,01
2	Гемоглобин	1 < 142,6	0,69
		2 142,6 - 148,0	1,34
		3 > 148,0	2,2
3	Средний объем эритроцита	1 < 75,133	0,61
		2 75,133 - 83,633	2,14
		3 > 83,633	0,74
4	Среднее содержание гемоглобина в эритроците	1 < 25,533	2,85
		2 25,533 - 28,893	1,62
		3 > 28,893	0,31
5	Средняя концентрация гемоглобина в эритроцитной массе	1 < 292	0,3
		2 292 - 344	1,36
		3 > 344	0,2
6	Распределение тромбоцитов по их объему	1 < 12,467	0,37
		2 12,467 - 13,4	1,74
		3 > 13,4	0,36
7	Лимфоциты %	1 < 26,88	0,01
		2 26,88 - 36,987	2,11
		3 > 36,987	0,69
8	Гранулоциты %	1 < 53,74	0,81
		2 53,74 - 58,9	2,86
		3 > 58,9	0,95

выступлений. В неё была внесена информация об используемых признаках и классах задачи, а также в базу фактов - НФД, представленного в виде порядкового номера, вектора признаков и известного номера класса. После анализа на непротиворечивость эти данные были транспортированы в базу знаний подсистемы.

На данном этапе экспертные знания в систему прогноза не вносились, т.к. не было экспертов по этой задаче для экспликации их суждений в базу знаний. Таким образом, база знаний содержала набор правил (продукций), основанных на фактических данных, а также имела в своем составе расчетные вероятностные оценки признаков и их градаций, включая интегральный показатель значимости признаков и значимость (влияние) каждой градации признака на результат успешности/неуспешности выступления спортсмена.

В результате была сформирована система прогноза успешности выступления пловцов на соревнованиях на основе гематологических показателей. Пример работы системы прогноза, реализованной в среде программирования Delphi на языке Object Pascal представлен на рис.2.

Для оценки качества прогноза была разработана специальная методика, основанная на методах статистического моделирования Монте-Карло, и соответствующее программное обеспечение. Методика заключается в следующем:

1. Все имеющиеся фактические данные в базе знаний системы (общая выборка) делятся случайным образом (с помощью генератора случайных чисел) на две: обучающую и контрольную. Контрольная выборка определяется (задается) процентом от общей выборки и вводится при настройке модуля оценки качества подсистемы в качестве параметра; Контрольная выборка содержит известный класс решений – успешный или неуспешный.

2. Обучающая часть выборки – вся оставшаяся часть общей выборки за исключением контрольной выборки, используется для формирования необходимых данных для трех алгоритмов прогноза.

3. Каждый случай из случайной контрольной выборки в цикле проходит автоматическую классификацию и сравнивается с его исходным (известным) классом;

4. Результаты расхождений проведенных циклов проверки суммируются и сохраняются;

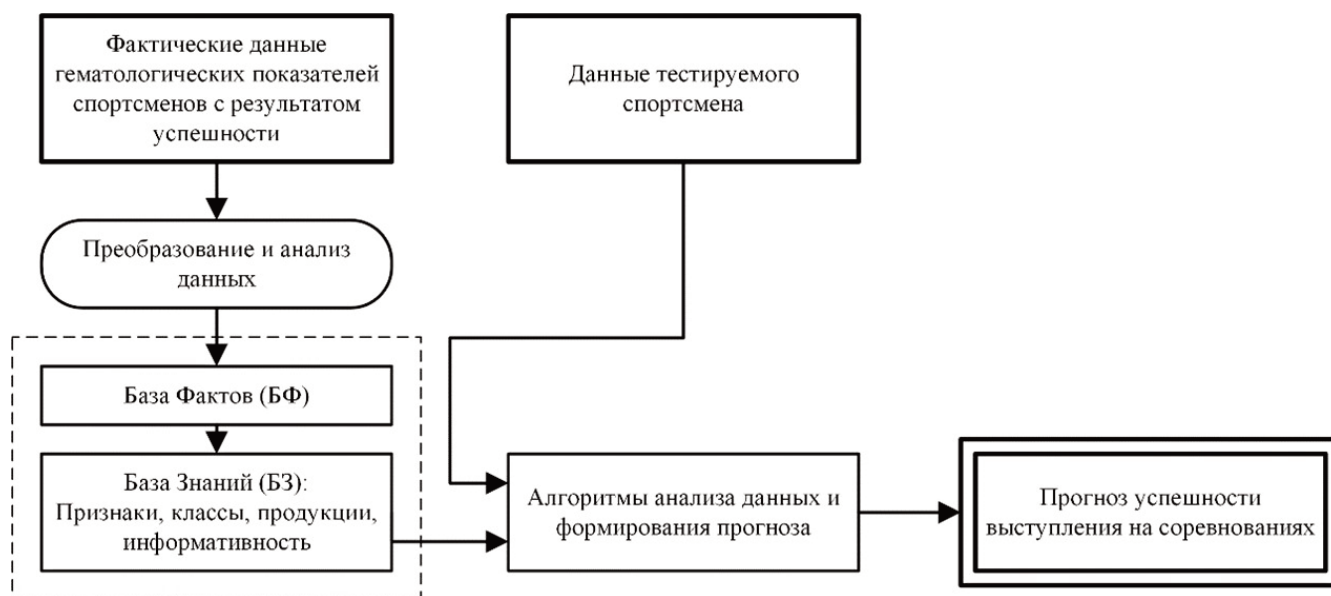


Рис. 1. Обобщенная схема системы прогноза успешности по гематологическим показателям пловцов

Fig. 1. Generalized diagram of a system for predicting success based on hematological indicators of swimmers

5. Такая процедура осуществляется несколько сотен раз, после чего вычисляются средние оценки расхождений (точность прогноза).

Предложенный подход обеспечивает статистически достоверные результаты оценки точности прогноза на основе небольшой случайной контрольной выборки за счет многократного повторения процесса прогноза. При этом нет необходимости отдельно иметь обучающую и контрольную выборки, что позволяет максимально использовать собранные фактические данные для построения систем классификации (прогноза). Точность прогноза оказалась около 93%, (при заданном объеме контрольной выборки в 10%), что позволяет сделать вывод о целесообразности использования такой подсистемы для формирования экспресс оценки прогноза успешности выступления спортсменов на соревнованиях.

Блок схема оценки точности методом Монте-Карло приведена на рис.3.

Выводы

Применение оценок интегральной значимости признаков позволило уменьшить число гематологических показателей крови с 18 до 8 для разработки системы экспресс прогноза успешности выступления спортсменов на соревнованиях.

Разработанная методика рационального разделения непрерывных признаков в дискретную шкалу

с темя градациями на основе критерия максимума интегральной значимости признака позволила дать простую и понятную для специалистов интерпретацию влияния градаций показателей на оценку прогноза результатов соревнований спортсменов пловцов.

Компьютерная система прогноза успешности выступления пловцов (мужчин) по результатам гематологических показателей спортсменов, полученных в соревновательный период, позволяет оперативно получить результат прогноза по каждому спортсмену и использовать его для отбора лучших кандидатов на этих соревнованиях.

Методика оценки качества заключений системы прогноза успешности выступления спортсменов на основе методов Монте-Карло не требует выделения отдельной контрольной выборки, использует для этого весь объем фактических данных и обеспечивает статистически достоверные результаты оценки точности прогноза.

Результаты статистического тестирования на подходах Монте-Карло показали достаточно высокую точность прогнозирования (около 93%) системы прогноза на имеющемся объеме фактических данных, что позволяет сделать вывод о ее работоспособности и возможности дальнейшего развития.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Финансирование. Исследование не имело спонсорской поддержки.

Участие авторов. Статья подготовлена с равным участием авторов.

Поступила: 15.09.2023. **Принята к публикации:** 16.10.2023.

Conflict of interest. The authors declare no conflict of interest.

Financing. The study had no sponsorship.

Contribution. Article was prepared with equal participation of the authors.

Article received: 15.09.2023. **Accepted for publication:** 16.10.2023

МКЛАСС Прогноз успешн.выступ.спортсмен./ГЕМАТОЛОГ

Параметры Заполнение БЗ РП Структ_зн Диагностика Дополн.

Вид Опроса: 0 - По порядку заведения признаков

Начать классификацию

Вектор Признаков: 23222222

☐ Классификация автоматически

Признаки: ☒ Автоматически

1 - Лейкоциты

Градации признака: 2 - 3,587 - 6,708

Классификация

Процедуры Похожие

Описание случая и результата классификации

1 - Лейкоциты: 3,587 - 6,708
 2 - Гемоглобин: > 148,000
 3 - Средний объем эритроцита: 75,133 - 83,633
 4 - Среднее содержание гемоглобина в эритроците: 25,533 - 28,893
 5 - Средн.концентр-я гемоглобина в эритроцитной массе: 292 - 344
 6 - Распределение тромбоцитов по их объему: 12,467 - 13,400
 7 - Лимфоциты %: 26,880 - 36,987
 8 - Гранулоциты %: 53,740 - 58,900

Результат Классификации "К-близ.соседей":
 * Имеется полное сходство по одному классу решений *
 Класс: 2 - Успешный

Оценка по Вероятностному алгоритму
 Вер-ть Класс Наименование
 99,9 2 - Успешный

Сходство по оценкам значимости признаков
 Сход-во Класс Наименование
 37 1 - Неуспешный
 63 2 - Успешный

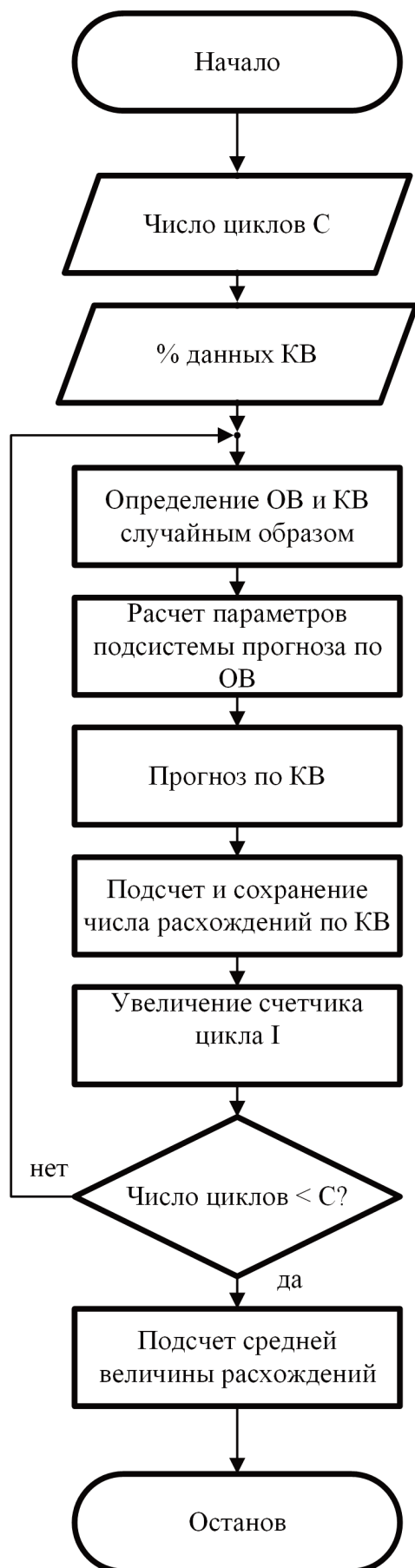
*** Комитетное заключение: ***
 Заключение основано на полном сходстве
 Класс: 2 - Успешный

КЛАСС	НАИМ_КЛ	КОЛ_ВО
2	Успешный	1

Рис. 2. Вид экрана заключения системы прогноза по гематологическим показателям спортсменов
 Fig. 2. View of the conclusion screen of the prognosis system for hematological indicators of athletes

СПИСОК ИСТОЧНИКОВ

- Сальников В.В., Болотов А.А., Барчуков В.Г., Тен А.М., Онощенко О.В. Ступенчатый тест для оценки функционального состояния пловцов в системе интеллектуальной поддержки тренера //Сборник материалов IV-й научно-практической конференции «День спортивной информатики». М., 2020. С. 119–123.
- Нехвядович А. И., Будко А. Н. Оценка эффективности тренировочного процесса спортсменов на основе вариабельности показателей крови: Практическое пособие. - Минск: БГУФК, 2019. 40 с.
- Суздальницкий Р.С., Левандо В.А. Иммунологические аспекты спортивной деятельности человека //Теория и практика физической культуры. 1998. № 10. С.43-46.
- Руководство по обслуживанию. Принципы работы. Описание электронного оборудования Chemwell-T, [Электронный ресурс]. URL: <https://www.manualslib.com/manual/1863455/Awareness-Technology-Chemwell-T-4600-Series.html?page=38#manual> (дата обращения: 15.01.2023).
- Таблица очков World Aquatics (FINA) [Электронный ресурс]. URL: <https://ruswimming.ru/fina-points> (дата обращения: 27.02.2023).
- Минцер О.П., Молотков В.Н., Угаров Б.Н., Попов А.А., Палец Б.Л., Кондратьев А.И., Бабий Я.С., Блантер Б.И. Биологическая и медицинская кибернетика / Справочник. - Киев: Наукова Думка, 1986. 375 с.
- Вопросы кибернетики. Задачи медицинской диагностики и прогнозирования с точки зрения математика. /под редакцией И.М.Гельфанда. - М.: АН СССР, 1985. 194 с.
- Ларичев О.И. О возможностях человека в задачах принятия индивидуальных решений при многих критериях. //Проблемы и методы принятия решений в организационных системах управления: Тр.конф. - М.: ВНИИСИ, 1982. С. 5-12.
- Ларичев О.И. Проблемы взаимодействия человек-ЭВМ в системах поддержки принятия решений. //Сб.трудов ВНИИСИ, - М.: ВНИИСИ, 1984 выпуск 9. С. 20-27.
- Ларичев О.И. Теория подсознательных решающих правил и ее применение в диагностических задачах // Психологический журнал. – 2003. – Т.24, № 1. С. 56-64.
- Simon H. Reason in human affairs. Stanford University Press, 1983. P. 115.
- Болотов А.А. Рациональные алгоритмы классификации в системе интеллектуальной поддержки тренера. // Сборник материалов III-й научно-практической конференции «День спортивной информатики». М., 2019. С. 132–139.
- Болотов А.А., Барчуков В.Г., Тен А.М., Онощенко О.В. Оценка информативности признаков в системе интеллектуальной поддержки тренера в циклических видах спорта. //Сборник материалов III-й научно-практической конференции «День спортивной информатики». М., 2019. С. 140–143.
- Болотов А.А., Барчуков В.Г. Методы и принципы построения полных и непротиворечивых баз знаний при дифференциальной диагностике двух близких по симптоматике заболеваний. //Клинический вестник ФМБЦ им. А.И. Бурназяна. 2022. № 4. С. 60–66.



REFERENCES

1. Salnikov V.V., Bolotov A.A., Barchukov V.G., Ten A.M., Onpchenko O.V. Step test to assess the functional state of swimmers in the system of intellectual support of the coach. Collection of materials of the IV-th scientific-practical conference "Day of sports informatics", Moscow, 2020, p. 119–123.
2. Nehvydovich A.I., Budko A.N. ASSESSMENT OF THE EFFICIENCY OF THE TRAINING PROCESS OF ATHLETES ON THE BASIS OF THE VARIABILITY OF BLOOD INDICATORS: A Practical Guid- Minsk: BGUFK, 2019. 40 p.
3. Suzdalskiy R.C., Levando V.A. Immunological aspects of human sports activity. Theory and practice of physical culture. 1998. № 10. p.43-46. (In Russian.)
4. Руководство по обслуживанию. Принципы работы. Описание электронного оборудования Chemwell-T, [Электронный ресурс]. URL: <https://www.manualslib.com/manual/1863455/Awareness-Technology-Chemwell-T-4600-Series.html?page=38#manual> (дата обращения: 15.01.2023).
5. Таблица очков World Aquatics (FINA) [Электронный ресурс]. URL: <https://russwimming.ru/fina-points> (дата обращения: 27.02.2023).
6. Mincer O.P., Molotkov V.N., Ugarov B.N., Popov A.A., Palec B.L., Kondratiev A.I., Babiy Y.C., Blanter B.I. Biological and Medical Cybernetics / Handbook, Kyiv: Naukova Dumka Publ, 1986, 375 p.
7. Gelfand I.M. Questions of cybernetics. Tasks of medical diagnostics and forecasting from the point of view of mathematician. Moscow, USSR Academy of Sciences Publ., 1985, 194 p.
8. Larichev O.I. On the possibilities of a person in the problems of making individual decisions under many criteria. Problems and methods of decision mking in organizational management systems. Moscow, 1982, p. 5-12.
9. Larichev O.I. Problems of human-computer interaction in decision support systems. VNIISI, v 9, Moscow, 1984, p. 20-27.
10. Larichev O.I. The theory of subconscious decision rules and its application in diagnostic problems. Psychological journal, 2003. Vol.24, № 1, p. 56-64.
11. Simon H. Reason in human affairs. Stanford University Press, 1983, 115 p.
12. Bolotov A.A. Rational classification algorithms in the coach's intellectual support system. Collection of materials of the III scientific-practical conference «Day of Sports Informatics». Moscow, 2019, p. 132–139.
13. Salnikov V.V., Bolotov A.A., Barchukov V.G., Ten A.M., Onpchenko O.V. Using probabilistic estimates of the significance of features to form recommendations and explain conclusions in the intellectual support system. Collection of materials of the III scientific-practical conference «Day of Sports Informatics». Moscow, 2019, p. 140–143.

Рис. 3. Блок схема алгоритма оценки точности прогноза методом Монте-Карло

Fig. 3. Flowchart of the algorithm for assessing the accuracy of the forecast using the Monte-Carlo method