

УДК 004

DOI: 10.46548/21vek-2020-0950-0021

РАЗРАБОТКА ВЕСОВОГО АЛГОРИТМА ДЛЯ ПРОГНОЗА ФУТБОЛЬНЫХ МАТЧЕЙ

©2020

Журавлев Александр Александрович, студент,

Рубин Олег Ильич, студент,

Аксенов Константин Александрович, кандидат технических наук, доцент,

Уральский Федеральный Университет им. Б. Н. Ельцина, ИРИТ РмФ

(г. Екатеринбург, ул. Мира, д. 32,

e-mails: SanyaProgrammer2503@gmail.com, olegrubin97@gmail.com, k.a.aksenov@urfu.ru)

Аннотация. В последнее время повысился интерес к прогнозированию событий в самых разных социальных областях. Очевидно, что для составления реалистического прогноза необходимо располагать специализированным алгоритмом, учитывающим множество влияющих факторов. В данной статье предложен весовой алгоритм прогнозирования результатов футбольных матчей, по которому для численной оценки ожидаемого исхода матча определенным факторам присваиваются весовые значения. Основными критериями для прогнозирования тех или иных показателей выбраны статус команды (домашний или гостевой), а также давность матча по отношению к прогнозируемому (чем раньше состоялся матч, тем меньший весовой коэффициент ему присваивается). Для анализа показателя берутся последние пятнадцать матчей команд, участвующих в прогнозируемом матче. Если текущий статус команды домашний, то матчи, в которых данная команда была домашней, имеют весовое значение 0,7, а гостевые – 0,3. Аналогичные показатели применяются и к гостевой команде (прошедшие гостевые матчи имеют весовой коэффициент 0,7, а домашние – 0,3). Давность матчей разделена на три категории: недавний (весовой коэффициент равен 0,5), матчи средней давности (весовой коэффициент – 0,3) и поздний (весовой коэффициент – 0,2). Показана работа данного алгоритма на конкретном матче. В качестве анализируемого показателя выбрано количество голов, которые могут забить команды. Анализ возможного количества забитых голов каждой из команд проведен с помощью присвоения весовых коэффициентов текущему статусу команд и давности анализируемых матчей. Получены данные, которые соответствуют реальности. Для оценки точности предложенного метода проведен анализ тысячи прошедших матчей. Сделаны выводы о корректности работы описанного метода. Корректным считается алгоритм, точность которого не меньше 80%. Исходя из полученного значения точности ($81,2\% \geq 80\%$), можно сделать вывод, что предложенный метод прогнозирования является корректным.

Ключевые слова: прогнозы, весовой алгоритм, футбольные матчи, точность алгоритма

DEVELOPMENT OF WEIGHT ALGORITHM FOR FORECAST OF FOOTBALL MATCHES

© 2020

Zhuravlev Alexander Alexandrovich, student,

Rubin Oleg Ilyich, student,

Aksenov Konstantin Aleksandrovich, candidate of Technical Sciences, Associate Professor,

Ural Federal University named after B. N. Yeltsin, IRIT RTF

(Yekaterinburg, Mira St., 32,

e-mails: SanyaProgrammer2503@gmail.com, olegrubin97@gmail.com, k.a.aksenov@urfu.ru)

Abstract. Recently, interest in forecasting events in various social fields has increased. Obviously, to make a realistic forecast it is necessary to have a specialized algorithm that takes into account many influential factors. This article proposes a weighted algorithm for predicting the results of football matches, according to which weighting values are assigned to certain factors to numerically evaluate the expected outcome of the match. The main criteria for predicting certain values are the status of the team (home or away), as well as the match age of the match relative to the predicted one (the earlier the match took place, the lower the weight coefficient is assigned to it). In the article, the main criteria for predicting certain values are the status of the team (home or away), as well as the prescription of the match (the earlier the match took place, the greater the weight coefficient) in relation to the predicted one. For the analysis of the value, the last fifteen matches of the teams participating in the predicted match are taken. If the current status of the team is home, then matches in which this team was home have a weighted coefficient of 0,7, and guest ones – 0,3. Similar values apply to the guest team (past guest matches have weight coefficient of 0,7, while home matches have weight of 0,3). The prescription of matches is divided into three categories: recent (weighting coefficient is 0,5), medium-old matches (weighting coefficient – 0,3) and late matches (weighting coefficient – 0,2). The work of this algorithm on a specific match is shown. As the analyzed value, the number of goals the teams can score is chosen. An analysis of the possible number of goals scored by each team was carried out by assigning weights to the current status of the teams and the prescription of the analyzed matches. Received data correspond to reality. To estimate the accuracy of the proposed method, an analysis of thousands of past matches was carried out. Conclusions are drawn about the correct work of the described method. An algorithm is considered correct if its accuracy is at least 80%. Based on the obtained accuracy value ($81,2\% \geq 80\%$), it can be concluded that the proposed forecasting method is correct.

Keywords: forecasts, weighting algorithm, football matches, algorithm accuracy.

Введение. Прогнозы являются очень популярным явлением. Предсказывать можно практически любые события: погоду, спортивные мероприятия, экономические показатели и пр. Для того чтобы иметь правильные прогнозы, необходимо наличие алгоритма, способного предоставлять корректные прогнозы на то или иное событие.

В данной статье описан алгоритм, позволяющий получить прогнозы на события, представляющие собой футбольные матчи. Для корректной работы метода необходимо вычислить его точность.

Целью данной статьи является разработка алгоритма для корректных прогнозов футбольных матчей.

В качестве **материала исследования** выступают матчи футбольных команд.

В статье используется эмпирический метод исследования, поскольку основным источником результатов являются сравнение и эксперимент.

Результаты и их обсуждение. В данной статье в качестве анализируемого события выбраны футбольные матчи, поскольку спортивные мероприятия пользуются большой популярностью.

Для предсказания корректных исходов на футбольные игры необходима разработка алгоритма, позволяющего осуществить данный процесс.

Алгоритм будет заключаться в следующем:

1. В качестве субъектов анализа выступают команды (домашняя и гостевая), проводящие матч.

2. Для анализа рассматриваются последние 15 матчей каждой из команд (домашней и гостевой).

3. Выбирается конкретный показатель для анализа (количество забитых голов, число желтых карточек или количество поданных углов той или иной командой).

4. В зависимости от временной близости последних игр к анализируемому матчу, присваиваются весовые коэффициенты: чем раньше матчи состоялись по отношению к анализируемой игре, тем больше весовой коэффициент имеет футбольная игра. В данной статье матчи будут разбиты на 3 категории: недавние матчи (последние 5 матчей команды по отношению к анализируемому матчу), матчи средней давности (последние 5 матчей по отношению к недавним матчам) и поздние матчи (5 матчей которые имеют более раннюю дату проведения по сравнению с другими матчами). Недавние матчи будут иметь весовой коэффициент 0,5, матчи средней давности – 0,3, а поздние матчи – 0,2 ($0,5 + 0,3 + 0,2 = 1,0$). Данные весовые коэффициенты выбраны на основе эксперимента, проведенного в разделе «Оценка точности весового алгоритма».

5. Также большое влияние играет статус команды в текущем матче (домашняя или гостевая команда), поэтому матчи, в которых статус команды был такой же, как и в анализируемой игре, имеют больший весовой коэффициент. Для домашней команды в анализируемом матче больший вес имеют матчи, в которых статус команды был домашним, а для

гостевой команды – гостевым. Пусть для домашней команды весовой коэффициент прошедших домашних матчей равен 0,7, а гостевых – 0,3. А для гостевой команды анализируемого матча, наоборот, прошедшие гостевые матчи имеют вес 0,7, а домашние – 0,3. Данные весовые коэффициенты выбраны на основе эксперимента, проведенного в разделе «Оценка точности весового алгоритма».

6. Выбранный показатель рассчитывается по формуле (1), которая представлена ниже.

Назовем выше описанный метод Весовым алгоритмом прогнозирования футбольных матчей.

Блок-схема описанного алгоритма представлена на рисунке 1.

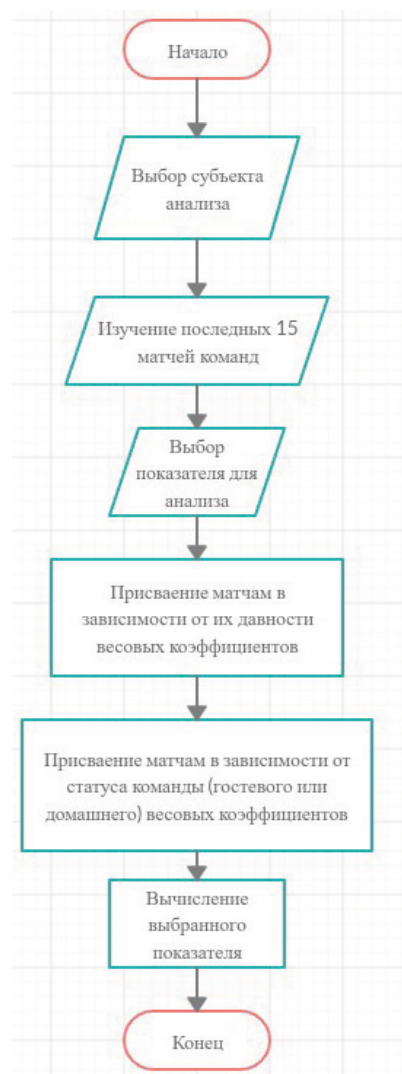


Рисунок 1 – Блок-схема Весового алгоритма

Введем формулу, позволяющую вычислить тот или иной показатель (количество ударов по воротам соперника, забитые голы и пр.) футбольного матча:

$$PM = BK_{ДС} * (BK_{НМ} * CG_{НДМ} / K_{НДМ} + BK_{СМ} * CG_{СДМ} / K_{СДМ} + BK_{ПМ} * CG_{ПДМ} / K_{ПДМ}) + BK_{ГС} * (BK_{НМ} * CG_{НГМ} / K_{НГМ} + BK_{СМ} * CG_{СГМ} / K_{СГМ} + BK_{ПМ} * CG_{ПГМ} / K_{ПГМ}) \quad (1)$$

где PM – показатель матча, BK – весовой коэффициент, CG – сумма голов, K – количество матчей,

ДС – домашний статус ГС – гостевой статус, НМ – недавние матчи, НДМ – недавние домашние матчи, НГМ – недавние гостевые матчи, СМ – матчи средней давности, СДМ – домашние матчи средней давности, СГМ – гостевые матчи средней давности, ПМ – поздние матчи, ПДМ – поздние домашние матчи, ПГМ – поздние гостевые матчи.

Проверка работы алгоритма на анализе футбольного матча. Проведем анализ возможного количества забитых голов домашней и гостевой командами. Для изучения показателя используем матч Барселоны (домашняя команда) – Borussia Dortmund (гостевая команда), состоявшегося 27 ноября 2019 года и завершившегося со счетом 3-1 в пользу Барселоны. Для анализа используем последние 15 матчей каждой из команд. В таблицах 1 и 2 приведено количество мячей забитых каждой из команд за указанный промежуток времени (последние 15 матчей). Данные по матчам брались с официальных сайтов Барселоны [19] и Borussia Dortmund [20].

Таблица 1 – Количество голов, забитых домашней командой (Барселоной) в последних 15 матчах

Дата матча	Соперник	Статус команды в матче	Количество забитых мячей	Давность матча
23.11.2019	Леганес	Г	2	Недавний
09.11.2019	Сельта Виго	Д	4	Недавний
05.11.2019	Славия Прага	Д	0	Недавний
02.11.2019	Леванте	Г	1	Недавний
29.10.2019	Вальядолид	Д	5	Недавний
23.10.2019	Славия Прага	Г	2	Средняя давность
19.10.2019	Эйбар	Г	3	Средняя давность
06.10.2019	Севилья	Д	4	Средняя давность
02.10.2019	Интер Милан	Д	2	Средняя давность
28.09.2019	Хетафе	Г	2	Средняя давность
24.09.2019	Вильярреал	Д	2	Поздний
20.09.2019	Гранада	Г	0	Поздний
17.09.2019	Borussia Dortmund	Г	0	Поздний
14.09.2019	Валенсия	Д	5	Поздний
30.09.2019	Осасуна	Г	2	Поздний

Примечание: Д – рассматриваемая команда имеет статус домашней команды, Г – команда является гостевой.

Используя формулу (1), получаем следующее количество голов для каждой из команд:

КГД (количество голов домашней команды) = $0,7 \cdot (0,5 \cdot (4 + 0 + 5) / 3 + 0,3 \cdot (4 + 2) / 2 + 0,2 \cdot (2 + 5) / 2) + 0,3 \cdot (0,5 \cdot (2 + 1) / 2 + 0,3 \cdot (2 + 3 + 2) / 3 + 0,2 \cdot (0 + 0 + 2) / 3) = 2.645$.

КГГ (количество голов гостевой команды) = $0,3 \cdot (0,5 \cdot (3 + 3 + 3) / 3 + 0,3 \cdot (1 + 2) / 2 + 0,2 \cdot (0 + 4) / 2) + 0,7 \cdot (0,5 \cdot (0 + 0) / 2 + 0,3 \cdot (0 + 2 + 2) / 3 + 0,2 \cdot (2 + 1 + 3) / 3) = 1.505$.

Исходя из полученных данных, можно сделать вывод, что домашняя команда забьет не менее 2 голов (2 и более), а гостевая команда минимум 1 мяч.

Как видно, полученные результаты КГД и КГГ соответствует действительности, поскольку Барселона (домашняя команда) забила 3 мяча (что не меньше

2 голов), а Borussia Dortmund – 1 мяч (что не меньше 1 гола).

Таблица 2 – Количество голов, забитых гостевой командой (Borussia Dortmund) в последних 15 матчах

Дата матча	Соперник	Статус команды в матче	Количество забитых мячей	Давность матча
22.11.2019	Падерборн	Д	3	Недавний
09.11.2019	Бавария Мюнхен	Г	0	Недавний
05.11.2019	Интер Милан	Д	3	Недавний
02.11.2019	Вольфсбург	Д	3	Недавний
26.10.2019	Шальке	Г	0	Недавний
23.10.2019	Интер Милан	Г	0	Средняя давность
19.10.2019	Borussia Mönchengladbach	Д	1	Средняя давность
05.10.2019	Фрайбург	Г	2	Средняя давность
02.10.2019	Славия Прага	Г	2	Средняя давность
28.09.2019	Вердер	Д	2	Средняя давность
22.09.2019	Айнтрахт	Г	2	Поздний
17.09.2019	Барселона	Д	0	Поздний
14.09.2019	Байер	Д	4	Поздний
30.08.2019	Юнион Берлин	Г	1	Поздний
23.08.2019	Кельн	Г	3	Поздний

Оценка точности весового алгоритма прогнозирования футбольных матчей. Для проверки корректной работы алгоритма необходимо оценить его точность. А для этого необходимо оценить его работу на гораздо большем количестве матчей.

Условимся называть корректным такой алгоритм, точность которого не меньше 80%.

Для оценки точности алгоритма проведен анализ 1000 матчей с уже известными результатами. Оценивалось количество забитых голов. Пусть, полностью предсказанный матч имеет вес 1,0, частично спрогнозированный (предсказано верно количество голов только одной из команд) – 0,5, а полностью неугаданный – 0,0. Кроме этого, будем изменять весовые значения коэффициентов для давности матчей и статуса команды (домашняя и гостевая), чтобы показать, что присвоенные ранее весовые коэффициенты, выдают наилучшую точность прогноза. В таблице 3 представлены результаты анализа.

Точность алгоритма имеет следующую формулу:

$$TA = (BK1 \cdot ПП + BK2 \cdot ЧП + BK3 \cdot НП) / ОК \quad (2)$$

где TA – точность алгоритма, BK1 – весовой коэффициент полностью предсказанных матчей, ПП – количество предсказанных матчей, BK2 – весовой коэффициент частично предсказанных матчей, ЧП – число частично предсказанных матчей, BK3 – весовой коэффициент неудачно предсказанных матчей, НП – число неудачно предсказанных матчей, ОК – общее количество проанализированных матчей.

Получаем:

$$\begin{aligned} TA1 &= (1,0 * 673 + 0,5 * 278 + 0,0 * 49) / 1000 = 0,812; \\ TA2 &= (1,0 * 604 + 0,5 * 324 + 0,0 * 72) / 1000 = 0,766; \\ TA3 &= (1,0 * 612 + 0,5 * 328 + 0,0 * 60) / 1000 = 0,776; \\ TA4 &= (1,0 * 564 + 0,5 * 344 + 0,0 * 92) / 1000 = 0,736; \\ TA5 &= (1,0 * 640 + 0,5 * 298 + 0,0 * 62) / 1000 = 0,789. \end{aligned}$$

Наилучшим показателем обладает эксперимент с весовыми коэффициентами: $BK_{ДС} = 0,7$, $BK_{ТС} = 0,3$, $BK_{НМ} = 0,5$, $BK_{СМ} = 0,3$, $BK_{ПМ} = 0,2$. Данный эксперимент обладает точность 0,812, что больше чем 0,8 (80%). Поскольку точность алгоритма превышает 80%, то его можно считать корректным.

Таблица 3 – Оценка точности Весового алгоритма прогнозирования матчей

№	$BK_{ДС}$	$BK_{ТС}$	$BK_{НМ}$	$BK_{СМ}$	$BK_{ПМ}$	Полностью пред- сказанные матчи	Частично предска- занные матчи	Неудачно пред- сказанные матчи	Общее количество матчей
1	0,7	0,3	0,5	0,3	0,2	673	278	49	1000
2	0,8	0,2	0,5	0,3	0,2	604	324	72	1000
3	0,6	0,4	0,4	0,4	0,2	612	328	60	1000
4	0,5	0,5	0,5	0,4	0,1	564	344	92	1000
5	0,7	0,3	0,4	0,3	0,3	640	298	62	1000

Закключение. Для того чтобы иметь правильные прогнозы, необходимо наличие алгоритма, способного предоставлять корректные прогнозы на то или иное событие.

В данной статье предложен весовой алгоритм прогнозирования футбольных матчей, который основан на присвоение весовых значений тем или иным показателям, например статусу команды в матче (гостевая или домашняя команда).

Для демонстрации работы предложенного алгоритма проведен анализ конкретного матча, в котором в качестве анализируемого показателя взято возможное количество забитых голов каждой из команд. Результат анализа можно считать удовлетворительным, поскольку полученное количество забитых голов каждой из команд практически соответствует реальности.

Для корректности работы данного алгоритма проведен анализ 1000 матчей, в которых обоснован выбор весовых коэффициентов для давности матчей и статуса команды. Точность алгоритма для лучшего случая составила 0,812, что превышает 80%. Поэтому предложенный метод можно считать корректным.

СПИСОК ЛИТЕРАТУРЫ:

1. Клейнберг, Дж., Тардос Е. Алгоритмы: разработка и применение. Классика Computers Science / Пер. с англ. Е. Матвеева. — СПб.: Питер, 2016. — 800 с.
2. Бхаргава А.. Грожаем алгоритмы. Иллюстрированное пособие для программистов и любопытствующих. - СПб.: Питер, 2017. - 288 с.
3. Хайнеман, Джордж, Пояис, Гэри, Сеяков, Стэнли. Алгоритмы. Справочник с примерами на C, C++, Java и Python, 2-е изд.: Пер. с англ. — СПб.: ООО «Альфа-книга», 2017. — 432 с.
4. Седжвик, Роберт, Уэйн, Кевин. Алгоритмы на Java, 4-е изд.: Пер. с англ. — М.: ООО «И.Д. Вильямс», 2013. — 848 с.
5. Лафоре Р. Структуры данных и алгоритмы в Java. Классика Computers Science. 2-е изд. — СПб.: Питер, 2013. — 704 с.
6. Алгоритмы на C++/ Р. Седжвик — М.: Национальный Открытый Университет «ИНТУИТ», 2016. - 1001 с.
7. Паронджанов В. Д.. Учись писать, читать и понимать алгоритмы. Алгоритмы для правильного мышления. Основы оптимизации. — М.: ЛМК Пресс, 2012. — 520 с.

8. Дасгупта С. Алгоритмы / С. Дасгупта, Х. Пападимитриу, У. Вазири; Пер. с англ. под ред. А. Шеня. — М.: МЦНМО, 2014. — 320 с.

9. Основы теории прогнозирования [В Интернете] // URL: <https://forecasting.svetunkov.ru/etextbook/about/>

10. Статистика и анализ футбольных команд [В Интернете] // URL: <https://nb-bet.com/Teams>

11. Футбольная статистика [В Интернете] // URL: <https://allsportlinks.net/statistic>

12. Подборка алгоритмов, которые правят миром [В Интернете] // URL: <https://tproger.ru/translations/algorithms-ruling-world/>

13. Как лучше всего изучать алгоритмы [В Интернете] // URL: <https://tproger.ru/tag/algorithms/>

14. Какие алгоритмы должен знать уважающий себя программист [В Интернете] // URL: <https://tproger.ru/experts/7/>

15. Jeff Erickson. Algorithms [В Интернете] // URL: <http://jeffe.cs.illinois.edu/teaching/algorithms/>

16. Algorithms, 4th Edition [В Интернете] // URL: <https://algs4.cs.princeton.edu/home/>

17. Jason Brownlee. Clever Algorithms: Nature-Inspired Programming Recipes [В Интернете] // URL: <http://www.cleveralgorithms.com/nature-inspired/index.html>

18. Steven S. Skiena. The Algorithm Design Manual [В Интернете] // URL: <http://www8.cs.umu.se/kurser/TDBAfl/VT06/algorithms/BOOK/BOOK/BOOK.HTM>

19. FC Barcelona | Official website [В интернете] // URL: <https://www.fcbarcelona.com/en/>

20. Borussia Dortmund [В интернете] // URL: <https://www.bvb.de/>

Статья поступила в редакцию 23.04.2020

Статья принята к публикации 10.06.2020