

Прогнозирование исходов спортивных событий с помощью применения стандартных стратегий на примере хоккейных матчей Кондратенко И. Р.¹, Левиков И. В.², Мельников В. А.³

¹Кондратенко Игорь Ринатович / Kondratenko Igor Rinatovich – магистрант,
направление: фундаментальная информатика и информационные технологии;

²Левиков Иван Владимирович / Levikov Ivan Vladimirovich – магистрант,
направление: бизнес-информатика,
кафедра информационных технологий и экономической информатики;

³Мельников Виталий Андреевич / Melnikov Vitalij Andreevich - кандидат экономических наук, доцент,
кафедра информационных технологий и экономической информатики,
Институт информационных технологий,
Челябинский государственный университет, г. Челябинск

Аннотация: предсказание исхода спортивных событий очень популярно среди фанатов по всему миру. В настоящий момент есть несколько продуктов, в чьи задачи входит решение этой проблемы. В данной научной статье рассказывается об одной из таких систем, использующую статистические методы для предсказания результатов хоккейных матчей в НХЛ (национальная хоккейная лига). Система была апробирована на 639 играх регулярного сезона 2015/2016. Верхняя граница предсказания была 65,69 % правильно предсказанных матчей.

Abstract: predicting sport outcomes is very popular among fans around the world. Nowadays, there are some system, that try to solve this problem. This article presents one such system, which uses statistics methods in order to predict the result of hockey games in NHL (National Hockey League). The system was approved on 639 games from regular part of 2015/2016 season. Upper-bound of prediction was 65,69 % correctly predicted matches.

Ключевые слова: прогнозирование, хоккей, верхняя граница предсказания, выигрыш хозяев, больше забитых, меньше пропущенных, лучшая разница, наибольшее количество очков.

Keywords: prediction, hockey, upper-bound, Home-team, Max GF, Min GA, Best GDiff, Max Points.

В современном мире прогнозирование занимает одну из самых важных ниш. Множество людей пытаются угадать, что будет дальше, как их действия отразятся на грядущем. Для своих целей они используют множество инструментов, от оккультных до научно-обоснованных методов. Не обошли стороной и спортивные события. Множество людей стараются угадать исход спортивных событий, количество которых ежедневно исчисляются тысячами. Естественно, их результат неизвестен. Хотя многие мечтают об обратном. Помимо букмекерских контор и их клиентов, это также важно для тренеров, самих спортсменов, а также спонсоров.

На сегодняшний день существует множество различных подходов при прогнозировании результатов. Каждый из них чем-то схож с другими и по-своему уникален. В качестве примера спортивных событий были выбраны хоккейные матчи в силу однозначности исхода (отсутствие ничейного результата), которые и станут объектом исследования. На настоящее время создано немалое количество моделей, которые предсказывают результат в диапазоне от 40 % до 70 %. Однако в [1] говорится, что upper-bound (верхняя граница вероятности) составляет около 62 %. В свою очередь, существует достаточное количество статей, опровергающих данную цифру [2]. При этом максимальная граница попадания достигается за счет построения метамоделей анализа данных, составляющие которой также представляют собой сложные модели с большим количеством параметров. Такие модели обычно используют методы интеллектуального анализа данных, причем не только числовых, но и текстовых [3]. Помимо построения новых моделей, некоторые авторы применяют уже существующие модели ко многим видам спорта. Так, автор статьи [4] использует модель предсказания футбольных матчей относительно хоккейных матчей.

Каждый, кто пытается решить проблему верного предсказания событий, составляет подобные модели, но при этом забывают об оценке стандартных (типовых), наиболее очевидных стратегий. В данной статье проверяется гипотеза эффективности применения стандартных стратегий при прогнозировании исходов хоккейных матчей.

Прежде чем приступить к непосредственному прогнозированию, необходимо подготовить данные для их последующей обработки. В качестве статистики были взяты игры Национальной Хоккейной Лиги сезона 2015/2016 – первая половина из них в качестве обучающей выборки, вторая половина в качестве тестовой, параметрами, в которых были забитые и пропущенные шайбы. Эти данные представляют собой сводную таблицу по каждой из 30 команд, приведенную на дату проведения предсказываемого матча. Таким образом, обучающая выборка подвергалась изменениям с целью накопления более достоверных данных и повышению точности прогноза. В ходе проверки состоятельности гипотезы были рассмотрены 5 стратегий: «Выигрыш команды-хозяйки» (далее – «Home-team»), «Выигрыш команды с

большим количеством заброшенных шайб» (далее – «Max GF»), «Выигрыш команды с наименьшим количеством пропущенных шайб» (далее – «Min GA»), «Выигрыш команды с лучшей разницей забитых и пропущенных шайб» (далее – «Best GDiff») и «Выигрыш команды с наибольшим количеством очков» («Max Points»). Результаты применения вышеописанных стратегий представлены в таблице 1.

Таблица 1. Результат применения стратегий

Стратегия	% угаданных матчей
Home-team	51,7073
Max GF	53,9837
Min GA	56,2602
Best GDiff	55,9350
Max Points	55,2846

Как видно из таблицы, самой эффективной стратегией является выбор команды, которая меньше всего пропускает. Причем процент угаданных результатов говорит о целесообразности применения типовых стратегий при прогнозировании хоккейных матчей. Результаты исследования могут служить как окончательным инструментом для решения поставленной проблемы, так и базовым средством, способствующим построению модели с целью повышения данных показателей, вследствие чего можно сделать вывод, что граница «upper-bound», заявленная в [1], действительно может быть превышена.

Литература

1. *Weissbock J.* Forecasting Success in the National Hockey League using In-Game Statistics and Textual Data [Текст] / Joshua Weissbock – 2014.
2. *Dubbs A.* Statistics-free sports prediction [Текст] / Alexander Dubbs. // Stat.AP. – 2015, arXiv:1512.07208v1.
3. *Weissbock J., Inkpen D.* Combining textual pre-game reports and statistical data for predicting success in the National Hockey League [Текст] / Joshua Weissbock // 27th Canadian Conference on Artificial Intelligence. – 2014.
4. *Marek P, Šedivá B., and Āoupal T.* Modeling and prediction of ice hockey match results [Текст] / Patrice Marek // J. Quantitative Analysis in Sports. – 2014. - № 10 (3). – С. 357–365.