

**Возможный подход к задаче прогнозирования спортивных результатов
методами анализа данных**

Научный руководитель – Рыжов Александр Павлович

Журавлев Артем Дмитриевич

Аспирант

Московский государственный университет имени М.В.Ломоносова,
Механико-математический факультет, Кафедра математической теории
интеллектуальных систем, Москва, Россия

E-mail: dac1964@mail.ru

Современные методы машинного обучения (ассоциативные правила, деревья решения, нейронные сети и т. д.) используются во многих областях для решения проблем ассоциации, классификации, сегментации, диагностики и прогнозирования. Вполне логично, что эти алгоритмы находят применение в такой экстремальной человеческой деятельности, как спорт уровня высоких достижений. В работе [1] авторы описали подход к прогнозированию спортивных события на примере футбольных матчей. В итоге исследователи создали распределение описывающее футбольный матч и с помощью него получили "теоретический" перевес над букмемерскими конторами в 5.5%. К недостаткам можно отнести то множество допущении, при которых было проведено исследование, что в итоге не позволило применить в жизни созданный алгоритм. Так как данные исследования имеют огромную финансовую значимость для рынка спортивных ставок, то других известных результатов в последние годы не было представлено.

Принимая во внимание огромное количество исторических данных по хоккею, можно предложить подход к прогнозированию хоккейных матчей - машинное обучение. Параметры игроков и матча вместе с результатом могут составить обучающую выборку. Алгоритм машинного обучения с учителем может использовать эту выборку для построения функции предсказания результатов новых матчей.

Целью настоящей работы является рассмотрение одного из возможных способов формализации задачи прогнозирования спортивных результатов и построение классификатора, имеющего предсказательную силу выше чем у букмекерских контор для определенных видов ставок. В результате работы был собран набор данных, описывающих последние 6 сезонов игр регулярного сезона КХЛ. В качестве одной из метрик оценки работы классификатор рассматривался ROI (return on investment) и по этому показателю был достигнут результат 5.2%

Источники и литература

- 1) Lisandro Kaunitz, Shenjun Zhong , Javier Kreiner, "Beating the bookies with their own numbers - and how the online sports betting market is rigged"