

Министерство образования Республики Беларусь

Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.89

Сушков
Артур Андреевич

Моделирование оценки рейтинга игровой команды на основе
квалиметрического метода

АВТОРЕФЕРАТ

на соискание академической степени
магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Научный руководитель
Лапицкая Н. В.
к.т.н., доцент

Минск 2014

ВВЕДЕНИЕ

Одной из областей, требующей прогнозирования будущих событий является спортивное прогнозирование т. к. в последнее время спорт привлекает все большее количество людей, аккумулирует значительные финансовые, материальные и интеллектуальные ресурсы, и постепенно превращается в один из важных элементов экономики. Для международной букмекерской конторы WilliamHill 2012 год был вторым подряд, в котором рост доходов составил более чем 20%. Чистые поступления компании оказались на 28% выше, а операционная прибыль выросла на 17%. Прогнозирование результатов спортивных соревнований само по себе является важной задачей, составляющей основу букмекерского бизнеса. Кроме того, эта задача может служить хорошим полигоном для тестирования различных методов экстраполяции и прогнозирования результатов процессов в условиях ограниченного статистического материала при большом количестве влияющих факторов, некоторые из которых заранее неизвестны.

Модели и компьютерные программы предсказания результатов спортивных игр разрабатываются на протяжении последних десятилетий. Большинство из них используют методы описания неопределенности: регрессивный и авторегрессивный анализ, метод Байезиана в комбинации с цепями Маркова и методом Монте-Карло. Особенности таких моделей являются: достаточно высокая сложность, большое количество допущений, потребность в наличии большого массива статистических данных. Кроме того, эти модели не всегда легко интерпретировать. Существуют также модели, использующие нейронные сети для предсказания результатов. Для них необходимо иметь огромные массивы статистических данных.

После получения результатов любой из систем прогнозирования появляется следующая задача – максимально эффективное использование полученных данных.

Существует множество стратегий игры с букмекером, но успех на длительной дистанции приносят лишь те, которые построены на выявлении в линии букмекера завышенных коэффициентов. Это происходит в том случае, когда букмекер недооценил вероятность определенного исхода.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является моделирование оценки рейтинга игровой команды на основе квалиметрического метода. Данная работа была проделана для создания системы прогнозирования результатов спор-

тивных событий и анализа линий букмекера на основе результатов, формирования ставок, отталкиваясь от текущего состояния счета.

Было проведено исследование и анализ существующих методов оценки.

Для достижения поставленной цели необходимо решить следующие задачи:

- проанализировать методы прогнозирования;
- изучить способы хранения статистики;
- построить статистику на основе имеющихся данных;
- сформировать входные данные для выбранного метода оценки вероятности на основе статистики;
- реализовать программное средство для экспериментального исследования выбранного метода оценки вероятности;
- провести экспериментальные исследования разработанной системы.

Объект исследования – результат спортивного события, в котором принимают участие два соперника.

Предмет исследования – оценка вероятности различных результатов спортивных событий.

Основной гипотезой, положенной в основу диссертационной работы, является возможность оценить вероятность исхода спортивного события, используя для этого методы, целиком реализуемые с помощью программного средства и базирующиеся на открытых статистических данных.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры "Программное обеспечение информационных технологий" по теме "Разработать модели, методы, алгоритмы для оценки параметров, повышения надежности и качества функционирования аппаратно-программных средств систем и сетей сложной конфигурации и внедрить в современные обучающие комплексы" (ГБ № 11-2004, № ГР 20111065, научный руководитель НИР - В. В. Бахтизин).

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя Н. В. Лапицкой, заключается в формулировке целей и задач исследования.

Опубликованность результатов диссертации

Результаты моего исследования на данную тему были опубликованы в сборнике научных трудов по материалам Международной научно-практической конференции "Наука, образование, общество: тенденции и перспективы".

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников, списка публикаций автора. В первой главе представлено описание азартных игр как элемента экономики и произведен анализ букмекерской деятельности. Поставлена задача и сформированы показатели для оценки рейтинга команд. Вторая глава посвящена исследованию существующих методов решения поставленной задачи. В третьей главе описаны итоги экспериментальных исследований и предложены методы улучшения выбранных методов. Общий объем работы составляет 54 страниц, из которых основного текста – 50 страниц, 5 таблиц на 2 страницах, список использованных источников из 39 наименований на 2 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во введении определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В первой главе проведен анализ существующих методов прогнозирования и классификации. Были выявлены достоинства и недостатки существующих методов и оценен потенциал решения поставленной задачи с применением описанных методов.

На данный момент разработано множество методов для получения прогнозов. При этом разделяют их на две группы: экспертные и формализованные. Которые в свою очередь разбиты на группы. Среди экспертных методов выделяют:

- методы с персональной экспертной оценкой,
- методы с коллективной экспертной оценкой.

Среди формализованных методов выделяют:

- методы прогнозируемой экстраполяции,
- методы моделирования,
- экономико-математические методы.

На рисунке 1 приведена схема классификации методов прогнозирования.

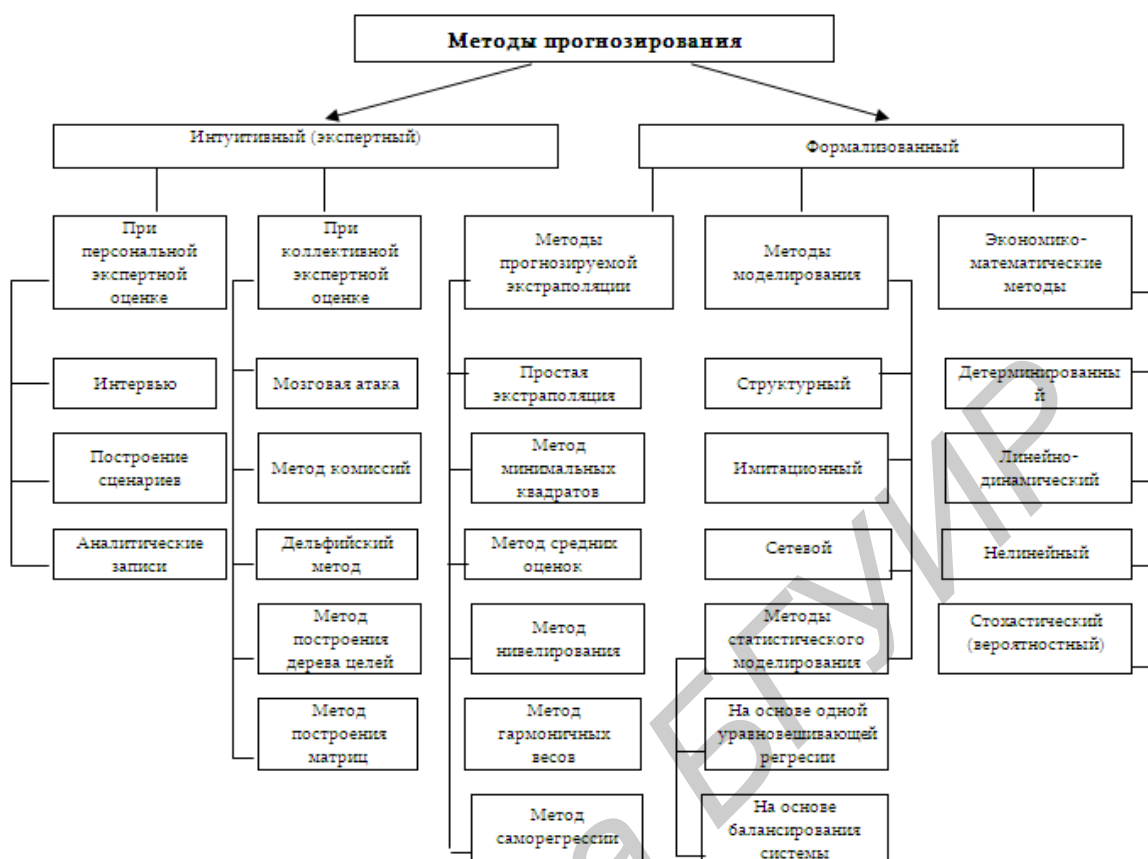


Рисунок 1 - Классификации методов прогнозирования

Результаты исследований, проведенных в этих направлениях, отражены в работах А.А. Грешилова, И.К. Бельской, Гвишиани Д.М., Веригин А.Н., Малютин С.А., Шашихин Е.Ю., Циндин Н. С., Акжигитова А. Н., Harrison P.J., Stevens C.F., Вороновский Г. К., Махотило К. В., Петрашев С. Н., Сергеев С. А. и др.

Экспертные методы были исключены по причине требования участия человека в процессе прогнозирования. Формализованные методы не применялись в сфере прогнозирования спортивных событий. В ходе экспериментов был испытан метод сумм и результаты оказались отрицательными.

В ходе анализа изученных методов был выявлен их основной недостаток для решения поставленной задачи – результат не несет вероятностный смысл и его сведение к вероятности не получается корректным.

Поэтому было принято решение изучить такую математическую модель как нейронные сети.

Во второй главе приводятся различные классификации нейронных сетей. Основаниями для деления на группы служат:

- характер обучения,
- настройка весов,
- тип входной информации,
- применяемая модель нейронной сети.

Классификация нейронных сетей представлена на рисунке 2.



Рисунок 2 – Классификация нейронных сетей

Среди применяемых моделей наиболее популярными являются сети прямого распространения. Их используют для распознавания, сжатия информации, классификации. Но их результат тоже не несет вероятностный характер, а значит, их использование не принесет наилучших результатов.

Рекуррентные сети не обладают стабильностью, что является критическим барьером для их применения на практике.

Наибольший интерес представляют сети радиально-базисных функций. Задачи классификации решает такая их разновидность как вероятностная нейронная сеть.

Пример архитектуры вероятностной нейронной сети представлен на рисунке 3.

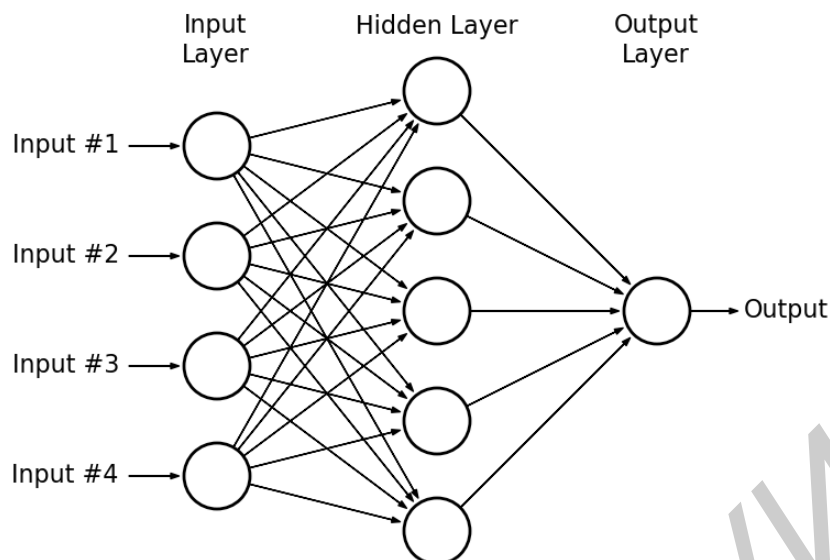


Рисунок 3 – Пример архитектуры вероятностной нейронной сети

Среди преимуществ вероятностной нейронной сети значится легкое обучение и реализация. Главные недостатки:

- узкая область применения,
- требует много места для хранения данных.

Но первый недостаток скорее является преимуществом т. к. поставлена именно задача классификации, а это является областью применения этих нейронных сетей. Так же для современных компьютерных систем проблема хранения данных легко решаема.

В итоге было принято решение выбрать в качестве модели прогнозирования и реализовать для экспериментального тестирования именно вероятностную нейронную сеть т. к. она более всего подходит для решения задачи классификации и результаты ее работы несут вероятностный смысл.

Третья глава описывает результаты экспериментального исследования выбранной модели прогнозирования. В ходе экспериментов были опробованы стандартные способы улучшения результата. Среди них уменьшение размерности входных данных и облегчение задачи классификации.

Уменьшение размерности входных данных было достигнуто благодаря анализу текущих входных данных. В ходе которого было выявлено то, что входные данные составляют три выборки, относящиеся к команде хозяйке, три выборки, относящиеся к команде гостю и две, относящиеся к ним обоим одновременно. Был сделан вывод, что выборки, относящиеся только к первой команде и только ко второй можно объединить между собой и получить вместо шести три выборки. Что уменьшит размерность входных данных на девять. Объединение выборок было осуществлено с помощью нахождения среднего арифметического между соответствующими статистическими вероятностями.

Облегчение задачи классификации было осуществлено благодаря разделению задачи классификации между тремя возможными исходами на три подзадачи – определение вероятности каждого из исходов независимо от двух других. Каждую из этих задач решала отдельно созданная и обученная нейронная сеть.

Так же был предложен новый для вероятностной нейронной сети способ нахождения вероятностей исходов – считать вероятность не в одной точке пространства, а в области вокруг нее, что уменьшит влияние локальных минимумов и максимумов.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Выбран метод оценки вероятности исходов спортивных событий. В ходе выполнения работы был проведён обзор литературы с целью анализа методов оценки вероятности. Были выявлены достоинства и недостатки различных методов и оценен потенциал решения поставленной задачи с их помощью. В итоге для получения результатов прогнозирования была выбрана вероятностная нейронная сеть. Произведены экспериментальные исследования выбранного метода, которые показали его эффективность и возможность применения для решения поставленной задачи.

2. Выбран эффективный метод хранения статистики. В ходе выполнения работы были изучены и проанализированы различные методы хранения данных и выбран наиболее подходящий. Для хранения статистики использовалась сериализация в файлы. Решение было принято, опираясь на размеры статистических данных и частоту операций чтения и записи, так же на удобство работы с данными.

3. Проанализированы показатели, характеризующие результаты спортивных событий и построены их оценки. Построена статистика на основе результатов спортивных событий, которые находятся в свободном доступе в сети интернет. Результаты состоят из названий команд соперников, даты игры и ее исхода. Статистика же отражает степень успешности каждой команды за любой период, а так же различные особенности этой команды, такие как умение пользоваться фактором домашнего поля или наоборот успешная игра в гостях, приверженность к открытой или закрытой игре, наличие удобных или неудобных соперников.

4. Определены и сформированы необходимые входные данные для оценки вероятности на основе статистики. В ходе формирования входных данных были учтены особенности выбранного метода оценки вероятности.

5. Спроектировано и реализовано программное средство для экспериментального исследования методов оценки вероятности. В программном средстве реализованы такие функции как получение нужных данных из открытых источников, получение статистики из данных и ее хранение, реали-

зованы выбранные методы оценки вероятности и нужные для них дополнительные функции, так же создана подсистема симуляции для проведения экспериментальных испытаний.

6. В ходе экспериментов были подробно изучены особенности выбранного метода оценки вероятности и выявлены скрытые недостатки. Были предложены методы их устранения и тем самым улучшения результатов.

Рекомендации по практическому использованию результатов

1. Полученные результаты формируют теоретическую и практическую базу для разработки ПО компьютерных систем для решения задач оценки вероятности исходов различных спортивных событий.

2. Разработанные методы и алгоритмы оценки вероятности спортивных событий могут применяться в автоматизированных системах букмекерских контор.

3. Результаты работы могут использоваться при подготовке персонала для разработки и обслуживания компьютерных систем, решающих задачи оценки вероятности любого рода событий.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

1-А. Матюшик В.Н., Сушков А.А. Методы и средства стеганографии для защиты графических образов. Статья подана для публикации в сборнике научных трудов по материалам «Международной научно-практической конференции». Часть III. 28 ноября 2014 г.

2-А. Сушков А.А. Практическое применение нейронных сетей. Статья подана для публикации в сборнике научных трудов по материалам «Международной научно-практической конференции». Часть III. 28 ноября 2014 г.

3-А. Сушков А.А. Вероятностные нейронные сети. Статья подана для публикации в сборнике научных трудов по материалам «Международной научно-практической конференции». Часть III. 28 ноября 2014 г.