

ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ АНАЛИЗА СОЦИАЛЬНЫХ И ЭКОЛОГИЧЕСКИХ ЯВЛЕНИЙ

Скулович О. З.

*Международный государственный экологический университет
им. А.Д.Сахарова, г. Минск*

В процессе изучения среды жизнедеятельности человека неизбежно происходит столкновение с огромным количеством воздействующих факторов различной природы, которые также могут давать многообразные результирующие эффекты в результате взаимного наложения и усиления. В самом широком смысле можно разделить среду обитания человека на природную, характеризующуюся непрерывностью своих подсистем, и прерывную - антропогенную. Антропогенная система в результате своего функционирования и развития оказывает увеличивающееся отрицательное воздействие на экологическую ситуацию внутри микросистем, ухудшая экологическую обстановку. Однако при разработке интегральной модели, описывающей среду жизнедеятельности человека, невозможно не учитывать воздействие социальной подсистемы. При таком многообразии факторов и зависимостей использование традиционных методов моделирования и прогнозирования может оказаться нерентабельным вследствие своей сложности, длительности разработки, а также невозможности в некоторых случаях справиться с задачами исследования[1].

Базируясь на вышеперечисленных аргументах, была поставлена задача изучения возможности применения нейросетевых технологий для решения задач анализа социальных и экологических явлений.

Исследуя возможность использования в качестве входных и выходных данных сети те или иные параметры, были приняты следующие решения. В качестве интегрального показателя, описывающего качество среды, и, соответственно, выходного параметра нейронной сети, было выбрано значение ожидаемой продолжительности жизни. Подробно о релеванности использования этой величины для описания качества жизни в системе «Человек – окружающая среда» (см. [2,3]). В качестве входных данных использовались статистические данные за последние 10 лет, описывающие различные факторы среды жизнедеятельности человека. В том числе социальные – ВВП, реальные располагаемые денежные доходы населения, показатель сельского хозяйства, показатель промышленности, уровень преступности, количество браков, разводов, уровень безработицы, уровень алкоголизма, численность врачей. А также следующие экологические факторы – выбросы вредных веществ, в том числе свинца и его соединений, сернистого ангидрида, окиси углерода, окислов азота, летучих органических соединений (ЛОС), углеводородов (без ЛОС), соединений серы (в пересчете на серу); уровни сброса сточных вод, образования токсичных отходов, мощности экспозиционной дозы излучения. На базе Neural network Toolbox пакета MatLab 7.0 была построена однонаправленная нейронная сеть прямой передачи сигнала.

Т.к. объем статистических данных оказался очень ограниченным, для проверки адекватности модели, помимо стандартных операций, исследовалось в том числе поведение сети в условиях с заранее предсказуемыми результатами. При прогнозе на будущее и постоянном увеличении объективно неблагоприятных факторов среды созданная нейронная сеть давала ожидаемый результат уменьшения ожидаемой продолжительности жизни. Аналогичный опыт, проводившийся для увеличения благоприятных факторов, также дал непротиворечивый результат.

Для этой же нейронной сети было проведено исследование наиболее значимых, с точки зрения влияния на продолжительность жизни, факторов, а также факторов, действие которых усиливается с течением времени [2,3]. В ходе анализа каждый фактор по очереди фиксировался на уровне 1995 г. и по величине отклонения результатов работы сети на этих данных и на всем массиве судилось о значимости этого фактора. Среди социальных факторов максимальное отклонение, а, следовательно, и степень влияния на

продолжительность жизни отмечается для ВВП и продукции сельского хозяйства. Кроме того, явно заметно наличие кумулятивного эффекта для таких факторов, как уровень ВВП и алкоголизм. Те же исследования для экологических факторов показывают максимальный уровень влияния для величин выбрасываемых вредных веществ, относительно высокий уровень для сернистого ангидрида, окиси углерода, окислов азота, углеводородов (без ЛОС), мощности экспозиционной дозы излучения, сброса сточных вод. Сильный кумулятивный эффект заметен для соединений серы.

В опытах, направленных на определение ситуаций, в которых сеть давала бы неприемлемый выход, был отмечен серьезный недостаток сети – ее крайняя неустойчивость. Данная нейросетевая модель, содержащая 21 нейрон только во входном слое, а также 2 скрытых слоя с нейронами, оказалась сильно подвержена изменениям даже при малых флуктуациях параметров.

Оптимизация нейросетевой модели проходила в двух направлениях – выделение среди всех входных параметров ведущих и несущественных факторов, а также уменьшение количества нейронов и слоев сети[2].

Была собрана база данных по шести областям Республики Беларусь и г. Минску, включающая 16 параметров, доступных в статистических справочниках регионов Республики. Благодаря расширению статистической базы, стало возможным проведение оценки значимости параметров методами пошагового регрессионного, а также корреляционного анализа.

В результате были выявлены ведущие признаки по каждой области, отдельно для мужского, женского и всего населения в целом. Проведение поиска наиболее значимых параметров по республике в каждый отдельно взятый год позволило как выявить признаки, характеризующиеся наиболее устойчивым влиянием, так и определить, что ведущие признаки в 90-е гг. XX в. и после 2000 г. разнятся. По результатам обработки полученного набора ведущих факторов, а также с применением Neural Network toolbox пакета Statistica, количество входных параметров нейронной сети было сокращено до следующих восьми: реальные располагаемые денежные доходы населения, производство промышленности, количество преступлений, количество разводов, алкоголизм, выбросы вредных веществ, выбросы сернистого ангидрида, сброса сточных вод.

Для того, чтобы избежать переобучения, а также для повышения устойчивости, количество внутренних слоев и нейронов в модели также было уменьшено. В результате, нейронная сеть имеет 8 нейронов во входном слое (по числу компонент входного вектора), один скрытый слой, содержащий 5 нейронов, и 1 нейрон в выходном слое (по числу компонент выходного вектора) [2].

Результаты проверки подтвердили большую устойчивость оптимизированной модели по сравнению с изначальной при моделировании и прогнозе, что иллюстрируют рис. 1 (а,б).

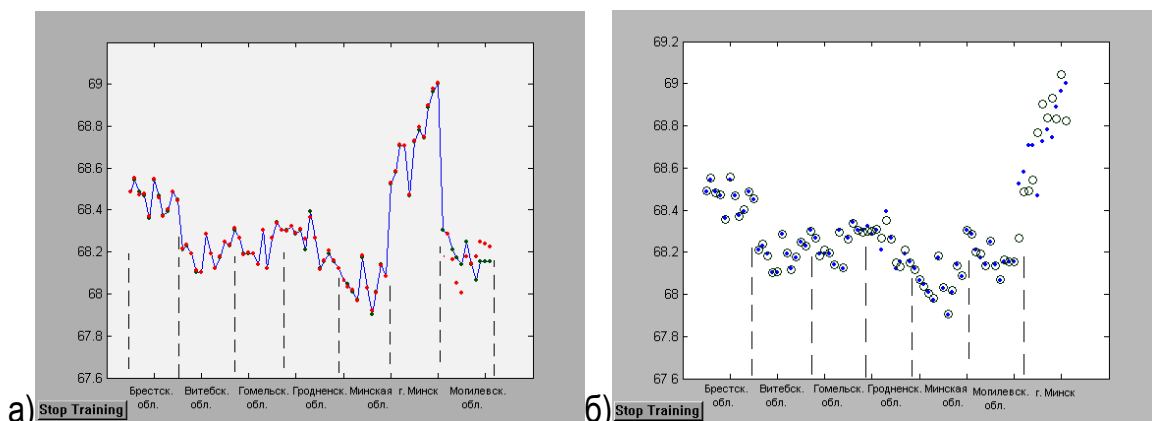


Рис. 1. Предсказание нейросетью величины ожидаемой продолжительности жизни для Могилевской области (а) и г. Минска (б)

Таким образом, не зависимо от того, лежат ли искомые значения прогноза в области средних по республике значений ОПЖ или значительно превосходят эти значения (как в случае прогноза ОПЖ для г. Минска), оптимизированная нейронная сеть справляется с прогнозом и показывает достаточную устойчивость.

В заключение стоит отметить, что оптимизация нейронной сети путем сокращения входных параметров до восьми ведущих (предварительно отобранных статистическими методами и специальными методами нейронных сетей) и упрощение ее внутренней структуры позволила значительно увеличить устойчивость сети при прогнозе.

Все эти тезисы применимы при разработке долгосрочных и краткосрочных программ по уменьшению негативного влияния окружающей среды на жизнь и здоровье человека, т. к. позволяют определить те приоритетные направления, на которые следует особо обратить внимание.

Литература

1. Тенденции заболеваемости, смертности и продолжительности жизни населения РБ. под общ. ред. д.с.н. Л.П. Шахотко – Минск, 2003.
2. Гринберг А.С., Иванюкович В.А., Скулович О.З. Нейросетевое прогнозирование и анализ ожидаемой продолжительности жизни. // Управление информационными ресурсами: материалы IV Международной научно-практической конференции. Минск, 17 мая 2006 г. – Мн., Акад. упр. при Президенте РБ, 2006.
3. Иванюкович В.А., Скулович О.З. Нейросетевая модель ожидаемой продолжительности жизни. / Тезисы докладов III международного экологического симпозиума в городе Полоцке: В 2-х т. Т. 2. – Полоцк: ПГУ, 2006.

МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОМОРФИЗМА НЕОРИЕНТИРОВАННЫХ ГРАФОВ

Теленкевич Р.С.

Брестский государственный технический университет, г. Брест

Введение

Важным разделом современной дискретной математики является теория графов. Родившись при решении головоломок и игр, эта теория стала мощным средством исследования и решения многих задач, возникающих при изучении больших и сложных систем.

Практическая роль графов возросла за последние годы в связи с интенсивным развитием таких направлений, как автоматизированные системы планирования, проектирования и управления, разработка и создание интеллектуальных систем обработки информации, представление задач, данных, знаний и методов поиска решений в искусственном интеллекте и т. д.

Одной из центральных комбинаторных проблем теории графов является проблема распознавания изоморфизма или эквивалентности двух графов. Она состоит в следующем: для заданных графов H и G требуется определить, существует ли взаимно однозначное отображение между множествами вершин графов, сохраняющее отношение смежности вершин.

Алгоритмы распознавания изоморфизма и изоморфного вложения графов необходимы при решении многих прикладных задач, например, таких, как:

- организация логического вывода на семантической сети;
- синтаксическое распознавание образов;
- эффективное представление в памяти ЭВМ структур данных;
- оптимизация вычислений в сети процессоров;
- автоматизация проектирования дискретных устройств ЭВМ и РЭА на основе унифицированного набора блоков;
- автоматизация контроля в САПР БИС;