

В.М. Беленький¹, Ю.В. Прус¹, В.Г. Спиридонов²
(¹Академия ГПС МЧС России, ²Ингосстрах; e-mail: av35740@akado.ru)

АЛГОРИТМЫ ФОРМИРОВАНИЯ БАЗЫ ДАННЫХ И ПОСТРОЕНИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПРОФЕССИОНАЛЬНОГО РИСКА НА ОСНОВЕ НЕЙРОПОДОБНЫХ СЕТЕЙ

Предлагается подход к моделированию профессионального риска на основе нейроподобных сетей. Описан программный модуль, обеспечивающий сопряжение базы данных с разработанной ранее программой "Нейросетевой предсказатель".

Ключевые слова: база данных, математическая модель, система управления, нейронная сеть, безопасность труда, регрессионный анализ.

V.M. Belenkiy, Yu.V. Prus, V.G. Spiridonov
**THE ALGORITHMS OF DATABASE GENERATION
AND BUILDING PROFESSIONAL RISK MATHEMATICAL MODELS
FOUNDED ON NEURAL NETWORKS**

The manner of professional risk modeling based on Neural networks is offered. The program module is described which provides the compatibility of database with early developed program "Network predictor".

Key words: data base, mathematical model, control system, neural network, labor safety, regression analysis.

Статья поступила в редакцию Интернет-журнала 17 июня 2015 г.

Построение моделей безопасности труда на основе нейроподобной сети

Для решения проблем обеспечения безопасности труда необходимо построение математической модели влияния факторов производственной среды на показатели профессионального риска работающих [1]. С этой целью может быть использован факторный анализ, методы нелинейной регрессии, а также различные варианты моделей, построенные на основе нейроподобных сетей [2]. Такие модели, в случае их достоверности, позволяют рассчитывать объективные прогнозы риска, оценивать реальный ущерб от неблагоприятных воздействий производства и планировать оптимальные комплексы профилактических мероприятий.

Одним из наиболее современных и достаточно точных методов моделирования является моделирование с помощью *нейроподобных сетей*. Идея нейронных сетей возникла в результате попыток смоделировать деятельность человеческого мозга. Мозг воспринимает воздействия, поступающие из внешней среды, и, используя память, обучается на собственном опыте. Модель нейронной сети состоит из простых нейронов, который объединены в группы (слои). У сети есть некоторое количество входов и выходов, которое определяется условием задачи.

Искусственная нейронная сеть представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Такие процессоры обычно довольно просты, особенно в сравнении с процессорами, используемыми в персональных компьютерах. Каждый процессор подобной сети функционирует только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам. И, тем не менее, будучи объединёнными в достаточно большую сеть с управляемым взаимодействием, такие локально простые процессоры способны выполнять довольно сложные задачи.

Построение модели нейронной сети осуществляется в несколько этапов:

- сбор данных для обучения;
- подготовка и нормализация данных;
- выбор топологии сети;
- экспериментальный подбор характеристик сети;
- экспериментальный подбор параметров обучения;
- собственно обучение;
- проверка адекватности обучения.

Нейронные сети особенно хорошо зарекомендовали себя при решении задач классификации, прогнозирования, кодирования и декодирования информации и, по сравнению с классическими математическими моделями, имеют ряд преимуществ. Например, при сравнении с регрессионной моделью, ошибка прогнозирования для нейронной сети значительно меньше:

$$E = \frac{1}{2} \sum_r \sum_j (Y'_{rj} - Y_{rj})^2,$$

где Y'_{rj} – моделируемое значение показателя риска Y таблицы исходных данных для слоя j и номера точки наблюдения r ;

Y_{rj} – реальное значение показателя.

В данном случае ошибка прогнозирования является ключевым показателем качества математической модели и её уменьшение даже на несколько процентов даёт существенное преимущество.

С использованием информации, обрабатываемой системой, пользователи в экстренных случаях проводят оперативную внеплановую работу либо ежемесячно составляют предложения по профилактике заболеваемости и охране труда для их включения в ежеквартальные и годовые планы мероприятий.

Для построения и апробации рассматриваемой модели нами проведены следующие исследования.

1. Разработана многофункциональная программа "Нейросетевой предсказатель", позволяющая получать прогноз показателей заболеваемости по заданным значениям факторов профессионального риска.

2. Разработан и реализован механизм загрузки в базу данных статистической информации по безопасности труда.

3. Реализован механизм отбора и выгрузки данных из базы в формате Excel.

4. Разработан алгоритм получения коэффициентов идентификации нейронных сетей (НС) с использованием программы "Statistica 6".

5. Реализован механизм ввода коэффициентов идентификации для обучения НС с использованием программы "Нейросетевой предсказатель".

6. Разработана даталогическая модель *базы данных (БД)*, содержащей коэффициенты идентификации НС, а также необходимую для моделирования информацию о заболеваемости и значениях факторов производственной среды. Таблицы БД и связи между ними реализованы с использованием реляционной СУБД MS Access 2007.

Для создания программы "Нейросетевой предсказатель" использована среда разработки "Delphi 7" (рис. 1).

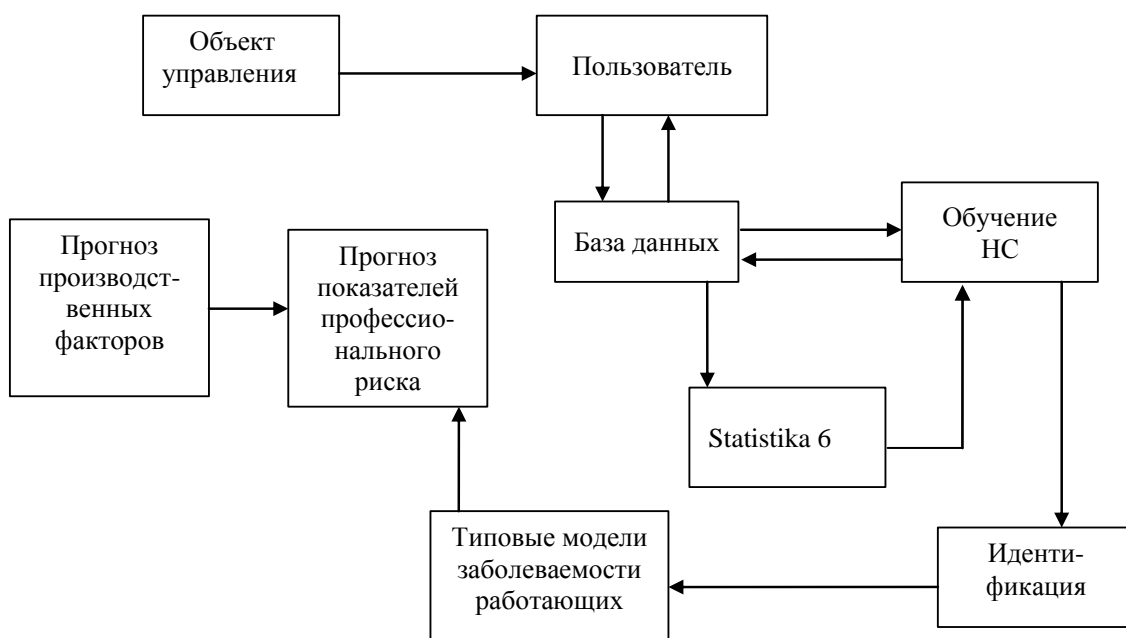


Рис. 1. Схема взаимодействия модулей программы "Нейросетевой предсказатель"

Модуль для работы с Excel-файлами в программе "Нейросетевой предсказатель"

Чтобы загрузить данные по безопасности труда в пакет Statistica, необходимо произвести их предварительную обработку. Это связано с тем, что для построения репрезентативной математической модели (как методами множественной регрессии, факторного анализа, так и на основе нейронной сети) нужно использовать сотни, а часто и тысячи точек наблюдения. При росте числа влияющих факторов объём исходных данных существенно возрастает.

Таким образом, возникает ряд задач, которые необходимо решить прежде, чем приступить к моделированию.

1. Для получения необходимых "точек наблюдения", как правило, требуется обработать тысячи "строк" первичных данных.

2. Первичные данные не всегда структурированы нужным образом, то есть имеют одинаковый формат, удобный для последующей обработки.

3. В первичных данных могут отсутствовать явные связи с остальными необходимыми для анализа и построения математической модели, данными (например, отсутствуют связывающие поля между данными по факторам условий труда и показателями заболеваемости).

Все эти задачи могут быть решены с использованием программы автоматизированной загрузки данных в базу, последующей их обработки и выгрузки в нужном формате.

Для этих целей в разработанную ранее программу прогнозирования "Нейросетевой предсказатель" встроен модуль, обеспечивающий загрузку, обработку и выгрузку данных по безопасности труда в формате Excel (рис. 2). При разработке модуля использовались данные РОССТАТ по аттестации рабочих мест и заболеваемости на предприятиях России. Суммарный объем исходных данных превышает 100 тыс. строк.

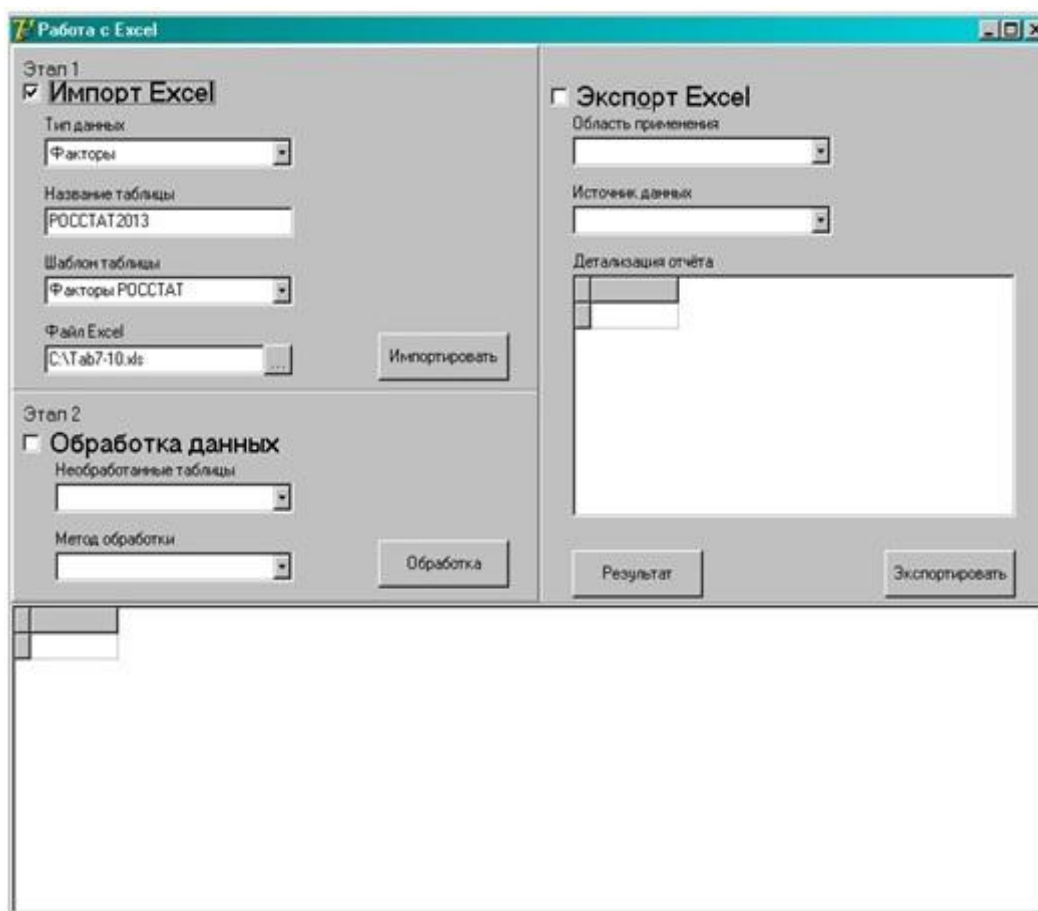


Рис. 2. Модуль работы с Excel-документами в программе "Нейросетевой предсказатель"

Загрузка данных разбита на два этапа. Первый этап содержит:

- создание новой таблицы в базе данных MS Access;
- первичную обработку исходной информации;
- загрузку в базу данных из Excel-файла.

Для реализации первого этапа через интерфейс программы задаётся ряд параметров.

Тип данных – тип загружаемых в базу данных. Для построения модели нейронной сети по безопасности труда необходимо 2 типа данных: факторы производственных воздействий и показатели заболеваемости.

Название таблицы – смысловая информация, определяющая характер загружаемых данных. На основании введённой информации, будет сформировано уникальное имя таблицы, в которую эти данные будут загружены.

Шаблон таблицы – тип структуры таблицы (наименование, тип данных и последовательность полей), в которую будут загружены данные. Структура таблицы хранится в самой программе "Нейросетевой предсказатель".

Файл Excel – путь к файлу в формате "excel", загружаемому в базу данных MSAccess.

На втором этапе происходит:

- обработка выгруженной в базу информации;
- перенос данных в нужную таблицу (группировка);
- задание параметров обработки данных.

Необработанные таблицы – список таблиц, которые были сформированы в базе, но ещё не обрабатывались.

Метод обработки – шаблонный алгоритм, по которому происходит обработка табличных данных. По окончании обработки все данные копируются в нужную таблицу (например, общую таблицу факторов условий труда или показателей заболеваемости).

В нижней сетке на форме работы с Excel-документами отображается пример данных, выгруженных в базу.

Экспорт информации из базы данных в Excel-файл аналогичен типовому механизму выгрузки отчёта. Для этого задаётся несколько значений.

Область применения – в каких целях будет использоваться результат выгрузки. Целью получения информации в формате Excel является формирование нужного представления данных для использования их в дальнейшем при построении математической модели профессионального риска на основе нейронной сети. В дальнейшем выгрузка может использоваться и для других целей, например, для построения отчётов по безопасности труда.

Источник данных – шаблон выгружаемых данных.

Пример получаемых данных можно посмотреть на сетке данных "Детализация отчёта" на форме, по нажатию кнопки "Результат".

Выводы

Благодаря созданному модулю для работы с Excel-файлами в программе "Нейросетевой предсказатель" стало возможным обрабатывать и анализировать огромные массивы исходных данных, что позволяет проводить исследования в сфере безопасности труда быстро и эффективно. За счёт значительного возрастания выборки данных, модели на основе нейронных сетей дают репрезентативные прогнозы показателей профессионального риска, которые могут быть использованы при обеспечении безопасности труда производственного персонала.

Литература

1. *Беленький В.М.* Разработка и внедрение экспертных систем на промышленных предприятиях. М.: изд-во УРАО, 2006.
2. *Беленький В.М., Спиридонов В.Г.* Использование нейронной сети для идентификации влияния условий труда на заболеваемость работающих. М.: ИПУ, 2012.
3. *Худсон Д.* Статистика для физиков. М.: Мир, 1970.
4. *Заенцев И.В.* Нейронные сети: основные модели. Воронеж: Воронежский Государственный университет, 1999. 76 с.