

В.М. Беленький, М.Ю. Прус, В.Г. Спиридонов
(Академия ГПС МЧС России; e-mail av35740@yandex.ru)

ОЦЕНКА БЕЗОПАСНОСТИ ТРУДА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ РЕГРЕССИОННОГО АНАЛИЗА И НЕЙРОННЫХ СЕТЕЙ

Предлагается метод оценки безопасности труда с учётом профессиональной заболеваемости в зависимости от вредных условий труда.

Ключевые слова: модель прогнозирования, база данных, нейронная сеть, охрана труда, профессиональный риск.

V.M. Belenkiy, M.Yu. Prus, V.G. Spiridonov

EVALUATION OF LABOR SAFETY USING METHODS OF REGRESSION ANALYSIS AND NEURAL NETWORKS

The method of the safety evaluation considering occupational disease depending on the unhealthy working conditions is offered.

Key words: forecasting model, database, neural network, labor safety, professional risk.

Статья поступила в редакцию Интернет-журнала 30 июня 2016 г.

Одним из самых сложных в охране труда является получение статистических данных по факторам воздействия и заболеваемости на объекте исследования. Это обусловлено неоднородностью условий труда на различных предприятиях, профессиональной спецификой, определяющей вредные и опасные производственные факторы и вызываемые ими травматизм и заболеваемость работающих. Для разработки необходимых профилактических мероприятий требуется периодическое проведение медицинских обследований и планомерный сбор данных по всем влияющим факторам. Получение достаточного объёма статистической информации является необходимым условием для выявления взаимосвязей и построения моделей идентификации [1].

1. Систематизация исходных данных по профессиональной заболеваемости для выявления статистических зависимостей

Для моделирования профессиональной заболеваемости работников предприятий были использованы данные РОССТАТа по профессиональной заболеваемости, травматизму и факторам условий труда за период 2007-2012 гг. в разрезе отраслей производства и регионов Российской Федерации. Значения показателей вредных производственных факторов (не менее 10) по точкам наблюдения в исходной выборке представлены в отчётах РОССТАТа по условиям труда за аналогичный период времени.

Вся исходная информация приведена к показателю на 100 работающих. Данные представлены по 194 структурным единицам отраслей производства и 94 структурным единицам регионов Российской Федерации. Таким образом, одна точка наблюдения и, соответственно, весь объём исходной выборки для моделирования показателя профессиональной заболеваемости формируется на основе однородных данных по отрасли, региону и году регистрации. Данные РОССТАТа получены в формате "excel" и загружены в *базу данных (БД)* с использованием модуля загрузки "excel-файлов". Для получения необходимых подмножеств данных используются SQL-запросы к БД. Затем информация выгружается в формат "xls", удобный для работы с интегрированным пакетом *STATISTICA*. По результату анализа общей выборки, выделены три отрасли производства с наибольшим количеством работников, имеющих профессиональную заболеваемость: "Добыча каменного угля, бурого угля и торфа", "Металлургическое производство", "Деятельность воздушного и космического транспорта".

Профессиональная заболеваемость по каждой отрасли идентифицируется сопоставлением:

- международного классификатора болезней;
- перечня опасных и вредных веществ;
- перечня производственных факторов, которые могут приводить к возникновению профессиональных заболеваний;
- перечней проводимых работ на данном производстве.

Для получения прогноза профессиональной заболеваемости работников предприятий по этим отраслям построим модели с использованием методов регрессионного анализа и *нейронной сети (НС)*. Проводится математическое моделирование профессиональной заболеваемости по отрасли "Добыча каменного угля, бурого угля и торфа".

Сравним результаты прогнозирования по двум достаточно адекватным для решаемой задачи методам идентификации профессионального риска [2].

2. Идентификация профессиональной заболеваемости на основе моделей нелинейного регрессионного анализа

Уравнение регрессии имеет вид:

$$Y_1 = 0,075 + 0,022X_4 + 0,022X_7^2.$$

Обозначение показателя – Y_1 (на 100 работающих).

Наиболее значимые факторы:

X_4 – повышенная запылённость воздуха рабочей зоны (на 100 работающих);

X_7 – повышенный уровень ионизирующего излучения (на 100 работающих).

Результаты, полученные с использованием данной нелинейной регрессионной модели, представим в табл. 1.

Таблица 1

Результаты, полученные по регрессионной модели

| $R = 0,54$ | $R^2 = 0,29$ | Корректированный $R^2 = 0,25$ |
|---|----------------------------|----------------------------------|
| $F = 7,964$ | $F_{\text{табл.}} = 2,390$ | $p = 0,001$ |
| Свободный член: 0,075 | | |
| Ст. ошибка: 0,17 | $t(39) = 0,44$ | $p = 0,66$ |
| Стандартная ошибка оценки показателя заболеваемости = 0 | | |

В табл. 1: R – коэффициент множественной корреляции; R^2 – коэффициент детерминации; F и $F_{\text{табл.}}$ – статистические критерии Фишера; t – критерий Стьюдента; p – доверительная вероятность.

Коэффициент множественной корреляции $R = 0,54$, коэффициент детерминации $R^2 = 0,29$ – это означает, что 29 % дисперсии показателя профессиональной заболеваемости отрасли "Добыча каменного угля, бурого угля и торфа" обусловлено изменением двух выбранных вредных производственных факторов.

Для критерия Фишера $F = 7,96$, вероятность нулевой гипотезы $p = 0,001$. Это меньше уровня значимости $\alpha = 0,05$ и позволяет отвергнуть нулевую гипотезу. В данном случае, под нулевой гипотезой понимается отсутствие взаимосвязи между исследуемыми переменными.

Рассчитаем минимальное значение показателя профессиональной заболеваемости отрасли РФ "Добыча каменного угля, бурого угля и торфа" при фиксированном значении показателя повышенной запыленности воздуха рабочей зоны.

Используем полученную модель нелинейной регрессии для расчёта минимальных значений ведущего фактора X_7 , варьируя фиксированными значениями второго по значимости фактора X_4 (рис.1).

Таким образом, нелинейные регрессионные зависимости иллюстрируют влияние двух наиболее значимых факторов условий труда на показатели профессиональной заболеваемости в отрасли "Добыча каменного угля, бурого угля и торфа".

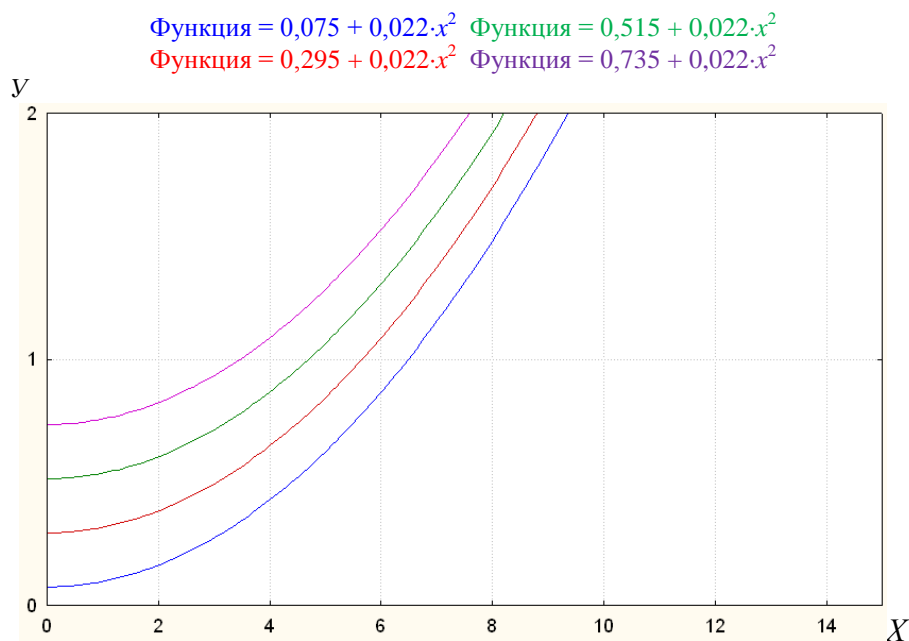


Рис. 1. Графики зависимости показателей профессиональной заболеваемости отрасли "Добыча каменного угля, бурого угля и торфа" (Y) от показателя ионизирующего излучения (X_7), при фиксированном значении показателя запылённости воздуха рабочей зоны (X_4)

3. Идентификация профессиональной заболеваемости на основе моделей нейронной сети

Значения тех же вредных производственных факторов для отрасли "Добыча каменного угля, бурого угля и торфа" задаём на входы разработанной ранее модели нейронной сети [3].

Задаём два входа модели НС:

- повышенный уровень запылённости воздуха рабочей зоны;
- повышенный уровень ионизирующего излучения.

На выходе модели получаем прогностические значения показателей профессиональной заболеваемости, соответствующие значениям вредных воздействий. Результат моделирования, с помощью мастера решений, представляет собой сотни вариантов различных нейронных сетей *многослойного перцептрона (МП)* и *радиальной базисной функции (РБФ)*, из которых нужно выбрать наиболее подходящие для исходной выборки данных (табл. 2).

Таблица 2

Основные характеристики выбранной сети

| Архитектура НС | Производительность обучения | Контрольная производительность |
|-----------------|-----------------------------|--------------------------------|
| РБФ 2-2-1-1-1 | 0,646 | 0,518 |
| Ошибка обучения | Контрольная ошибка | Тестовая ошибка |
| 1,120 | 1,487 | 3,825 |

Лучшие результаты показала нейронная сеть РБФ структуры 2-2-1-1-1 (рис. 2).

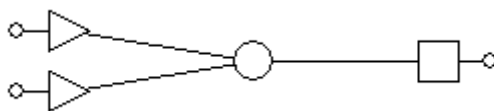


Рис. 2. Модель нейронной сети РБФ (2-2-1-1-1)

Сравним результаты полученных моделей *нелинейной регрессии (НР)* и нейронной сети (табл. 3). Для этого, вычислим прогноз показателей профессиональной заболеваемости Y_1 , полученный с использованием модели НС, вводя те же исходные значения, которые были заданы для модели регрессионного анализа.

Таблица 3

Сравнение результатов прогноза по моделям НР и НС

| X_4 | X_7 | Y_1 (НР) | Y_1 (НС) |
|-------|-------|------------|------------|
| 0 | 0 | 0,075 | 0,411 |
| 10 | 0 | 0,295 | 0,077 |
| 20 | 0 | 0,515 | 0,003 |
| 30 | 0 | 0,735 | 0,2 |

Выводы

1. Используя исходные статистические данные РОССТАТа по профессиональной заболеваемости и условиям труда для ряда отраслей производства Российской Федерации за 2007-2012 гг. построены модели регрессионного анализа и нейронной сети. В частности, для указанных моделей рассчитаны многофакторные зависимости показателей профессиональной заболеваемости и выявлены наиболее значимые вредные факторы условий труда по отрасли "Добыча каменного угля, бурого угля и торфа".

2. Сравнение результатов моделирования показывает, что предсказанные значения показателей профессиональной заболеваемости (Y_1) для моделей нейронной сети и регрессионного анализа различаются на всей выборке исходных данных, при этом порядок предсказанных значений по обеим моделям совпадает.

3. Таким образом, разработанные модели в сочетании с пакетом *STATISTIKA* позволяют с заданной доверительной вероятностью прогнозировать показатели профессионального риска в зависимости от уровня вредных факторов производственной среды.

4. На основании моделей создан программный продукт "Нейросетевой предсказатель", получивший Свидетельство о Государственной регистрации [4]. Данный программный модуль будет использован в автоматизированной системе управления безопасностью труда подразделений противопожарной службы, разработка и апробация которой включена в план Академии ГПС МЧС России на 2016-2020 гг.

Литература

1. *Ким Дж.-О., Мьюллер Ч.У., Клекка У.Р. и др.* Факторный, дискриминантный и кластерный анализ: пер. с англ. / Под ред. И.С. Енюкова. М.: Финансы и статистика, 1989. 215 с.
2. *Прус Ю.В., Беленький В.М., Спиридонов В.Г.* Модели идентификации и эвристические алгоритмы принятия решений в системах управления безопасностью труда // Эвристические алгоритмы и распределённые вычисления. 2014. № 1. Вып. 6. С. 6-19.
3. *Заенцев И.В.* Нейронные сети: основные модели. Воронеж, 1999. 76 с.
4. *Беленький В.М., Спиридонов В.Г.* Прогнозирование показателей заболеваемости работающих с использованием обучаемой нейронной сети и интерфейсного модуля, разработанного в программной среде Delphi 7 // Нелинейный мир. 2014. № 4. Т. 12. С. 64-69.