УДК 378.675

С. А. Горбатков, д-р техн. наук, проф.,
проф. каф. математики и информатики, e-mail: sgorbatkov@mail.ru,
Финансовый университет при Правительстве Российской Федерации, Уфимский филиал, г. Уфа,
Д. В. Полупанов, канд. техн. наук, доц.,
доц. каф. информационных технологий и компьютерной математики, e-mail: demetrious@mail.ru,
Башкирский государственный университет, г. Уфа

Оптимальный отбор и агрегирование экзогенных переменных в нейросетевых моделях банкротств на основе функций Харрингтона

Предложены два метода сокращения размерности факторного пространства при построении нейросетевых моделей банкротств: оптимального отбора факторов и агрегирования с использованием обобщенной функции желательности Харрингтона. Общей оригинальной концепцией для обоих методов, отличающейся от известных методов компрессии переменных, является взаимосвязь с операцией регуляризации обратной задачи обучения нейросетей на байесовском ансамбле.

Ключевые слова: *оптимальный отбор, агрегирование факторов, нейросетевые модели банкротств, системный подход к моделированию*

Введение. Актуальность исследования

Работа посвящена одному из важных вопросов совершенствования нейросетевых технологий в сложных условиях моделирования (неточности, неполноты, неопределенности данных), характерному для задач банкротств. Широко известен факт, что предобработка данных и, в частности компрессия факторного пространства, предопределяет качество получаемой нейросетевой модели (НСМ), особенно в условиях сильной зашумленности данных и малого объема наблюдений [1]. Экономические приложения и более узко — модели банкротств не являются исключением [2—5]. Однако алгоритмы оптимального отбора факторов, а также компрессии факторного пространства в классе обратных задач для НСМ банкротств до сих пор не исследованы, точнее отсутствует общепринятый теоретический подход в этом вопросе. Исключение составляет работа Риссанена [6], где получен теоретически обоснованный обобщенный результат, базирующийся на принципе минимальной длины описания в выбранной модели-гипотезе h из ансамбля моделей, который выражается через информационную числовую меру Кульбака—Лейблера.

Подход Риссанена имеет существенное ограничение: он базируется на допущении об априорном знании закона распределения шумов в данных, который в экономических приложениях зачастую неизвестен. В настоящей работе мы отказались от этого допущения в целях приближения НСМ к практике.

Подчеркнем также важность системного подхода к разработке алгоритма предобработки данных. Если рассматриваемые НСМ восстановления зависимостей являются классом моделей, управляемых данными [7], то следует ожидать, что учет взаимосвязи алгоритмов предобработки данных (включая оптимальный отбор, компрессию факторного пространства) с алгоритмами обучения модели и ее регуляризации породит положительный эмерджентный эффект. Этот вопрос в литературе не рассматривался.

В литературе также не рассматривался вопрос об агрегировании переменных в НСМ на основе обобщенной функции желательности Харрингтона, что позволяет уже на стадии предобработки данных эффективно учитывать взаимодействие факторов, т. е. их влияние друг на друга.

Проведение исследований по указанным выше малоизученным актуальным вопросам применения нейросетей (НС) в моделях банкротств и послужило посылом к написанию данной статьи.

1. Постановка задачи исследования

Вначале оговорим экономическую постановку задачи, которая состоит в обслуживании банком своего кредитного портфеля в рамках выделенного примерно однородного кластера корпораций-заемщиков, сходных по отрасли экономики, условиям деятельности на рынке, экономико-политической обстановке в регионе и др. Фактор масштабности корпораций учитывается приведением экономических показателей стандартной бухгалтерской отчетности к безразмерному виду. Примером служат так называемые "финансовые коэффициенты" [3, 4], отражающие ликвидность корпораций, платежеспособность, рентабельность бизнеса, деловую активность и другие показатели финансово-экономического состояния. Число этих коэффициентов может доходить до нескольких десятков. Требуется построить модель для указанного кластера корпораций-заемщиков, которая позволяла бы оценивать вероятность банкротства конкретного заемщика и, главное, определять способы реструктуризации задолженности в зависимости от достигнутой стадии банкротства. При этом для построения модели предполагается использование ретроспективных данных аналогичного кластера заемщиков, содержащих для каждой корпорации информацию типа "банкрот — не банкрот". Для этих целей могут быть использованы базы данных самой кредитной организации, базы "Бюро кредитных историй Российской Федерации", налоговых органов.

Теперь изложим информационно-математическую постановку задачи исследования. Будем рассматривать обратную задачу (ОЗ) восстановления зависимости вероятности банкротства P от вектора экзогенных переменных $\mathbf{x} = (x_1, x_2, ..., x_j, ..., x_n)$, скрытой в данных. Эту зависимость будем определять в виде логистической функции, предложенной Ольсоном [8]:

$$P(t) = 1/[1 + \exp(-\hat{y}(\mathbf{x}(t), t))],\ P \in [0, 1], \quad (1)$$

где t — время, а показатель экспоненты $\hat{y}(\mathbf{x}(t), t)$, играющий роль аргумента в (1), восстанавливается с помощью НС-отображения:

$$\hat{y}(\mathbf{x}, t) = F(\mathbf{x}, \mathbf{W}, t),\ F\colon \mathbf{x} \in R^{(n)} \to \hat{y} \in R^{(1)}, \quad (2)$$

где $\mathbf{W}$ — множество синаптических весов НС; $F(\cdot)$ — оператор НС-отображения.

Отметим сразу особенности логистической НСМ (1)—(2).

1. Модель способна восстанавливать любые (даже очень сильные) нелинейные зависимости $y(\mathbf{x}, t)$, что выгодно отличает ее от регрессионных моделей Альтмана [3] и его последователей. В частности, в модели (1)—(2) исчезает проблема мульткоllинеарности факторов.

2. Логистическое отображение (1) является сжимающим в том смысле, что интервал для аргумента $\hat{y}(\mathbf{x}, t) \in [-7, 7]$ отображается в соответствующий интервал значения функции $P \in [0, 1]$, т. е. коэффициент сжатия составляет примерно 14. Следовательно, если НСМ (1)—(2) уже получена, протестирована и проэкзаменована, то ошибки в задании вектора факторов $\mathbf{x}$ будут "сжиматься" оператором логистического отображения (1) при вычислении вероятности P. Однако задача обучения НС, т. е. нахождения синаптических весов $\mathbf{W}$, является ОЗ, некорректно поставленной по Адамару [9], что требует специальных мер по регуляризации ее решения. Одним из путей является подход введения некорректной ОЗ в класс корректно поставленных по Тихонову [9]. Согласно теории решения некорректных задач, разработанной А. Н. Тихоновым и представителями его научной школы, показано: если в пространстве искомых решений Z выделить более узкое множество (компакт) $Z' \subset Z$, такое, что выполняются условия:

- *a priori* известно, что на компакте $\exists \mathbf{z} \in Z'$ (решение ОЗ существует);
- $\forall \mathbf{z}\ \ \exists \mathbf{z} \in Z' : \mathbf{z} = A^{-1}(\mathbf{u})$ (решение единственно);
- $\forall \delta\ \ \exists \xi : (\Delta \mathbf{u} < \delta) \Leftrightarrow (\Delta \mathbf{z} < \xi)$, если $(\mathbf{z} + \Delta \mathbf{z}) \in Z'$ (решение устойчиво),

то можно регуляризировать решение.

Подчеркнем основную идею теории А. Н. Тихонова: для регуляризации ОЗ следует сужать пространство Z искомых решений. В работе [9] разработаны такие конструктивные методы, в частности стабилизирующие функционалы $\Omega(\mathbf{z})$.

Применительно к рассматриваемой ОЗ восстановления параметров НСМ банкротств (1)—(2) роль искомой функции $\mathbf{z}$ играет множество параметров НС, т.е. синаптических весов $\mathbf{W}$, аналогом обратного оператора A^{-1} является обратный оператор НС-отображения F^{-1} из выражения (2). В качестве наблюдаемых экспериментальных характеристик объекта $\mathbf{u}$ выступают кортежи (вектор-строки) таблицы выборки данных $\langle \mathbf{x},\ y \rangle_{ig}$, $i = \overline{1, N}$, $g = \overline{1, G}$, где y_g — метки "банкрот — не банкрот" для корпорации-заемщика; g — номер объекта; i — номер наблюдения. В качестве стабилизирующего функционала Тихонова в нашей задаче может быть выбран $\Omega(\mathbf{W}) = \|\mathbf{W}\|^2_{E_n}$, где $\|\cdot\|_{E_n}$ — эвклидова норма матрицы $\mathbf{W}$. Такой пример рассмотрен авторами работы [4].

Однако даже при теоретической корректности ОЗ возможно появление практической некорректности. Причинами могут быть наличие шумов в экспериментальных данных (потеря устойчивости решения) и дискретность набора экспериментальных точек (неединственность решения) — наблюдаемые характеристики объекта могут быть описаны несколькими различными функциями $\mathbf{z}_1 = F_1^{-1}(\mathbf{x}, y)$, $\mathbf{z}_2 = F_2^{-1}(\mathbf{x}, y)$ и др. В случае НС $F_1(\cdot)$, $F_2(\cdot)$— это различные операторы НС отображения (2).

Так, в работе [3] авторами на модельном примере НС-аппроксимации искусственно зашумленной детерминированной зависимости было показано, что существует критический уровень зашумленности и ее объема в данных по доле зашумленных вектор-столбцов $\{x_{ij}\}$ и y_i, когда НС теряет устойчивость.

Следовательно, требуется специальный алгоритм регуляризации ОЗ обучения НС. Этот вопрос применительно к НСМ банкротств практически не изучен. Авторы статьи разработали на базе подхода С. А. Шумского [10] алгоритм регуляризации обучения НС, названный "квазибайесовским" (см. ниже в разд. 2 настоящей работы). Этот термин означает упомянутый выше во введении отказ от допущения об априорном задании вида закона распределения шумов в данных и использует парадигму А. Н. Тихонова сужения пространства искомых решений.

3. При формировании обучающего множества НС возникает проблема, которую мы назвали "динамической неполнотой данных". До исследования авторов [3] решение этой проблемы не рассматривали. Подчеркнем, что "динамическая неполнота данных" при обучении НС — это одно из наглядных свойств НСМ как модели, управляемой данными. Суть этой проблемы состоит в следующем.

Пусть имеются ретроспективные данные наблюдения входного вектора $\mathbf{x} = (x_1, \ldots, x_j, \ldots, x_n)$ в нескольких временных срезах $\{t_k\}$, $k = \overline{1, N}$. Обучающая выборка $\langle \mathbf{x}_{gk}, y_{ig} \rangle$, $g = \overline{1, G}$ содержит наблюдения для G предприятий. При этом в последнем временном срезе ($k = N$) известны как векторы значения факторов $\{\mathbf{x}_{gN}\}$, так и значения выходной переменной $y = \arg\left(P\left(\mathbf{x}_{gN}, t_N\right)\right)$. Это позволяет обучить НС и оценить вероятность риска банкротства P в последнем временном срезе ($t = t_N$). Однако для всех предшествующих временных срезов $(t_1, t_2, \ldots, t_{N-1})$ значения выходной (эндогенной) переменной $y_{gk} = \arg\left(P\left(\mathbf{x}_{gk}, t_k\right)\right)$, $k = \overline{1, N-1}$ могут быть неизвестны, поскольку к моменту $t \leqslant t_N$ для части корпораций-заемщиков процесс кризиса развивается и они еще не признаны банкротами. В этом и проявляется глубокая специфика моделируемых объектов. Нужно построить специальный итерационный процесс восстановления эндогенных переменных $\{y_{gk}\}$, $k = \overline{1, N-1}$, $g = \overline{1, G}$ во всех временных срезах и для всех объектов. Соответствующий нейросетевой логистический динамический метод (НЛДМ), базирующийся на системном законе инерционности экономических процессов, разработан авторами в работе [3]. Описание этого метода выходит за рамки статьи. Поэтому далее будем считать, что данные во всех временных срезах, т. е. кортежах $\langle \mathbf{x}_{gk}, y_{gk}, t_k \rangle$, восстановлены и являются полными и НСМ рассматривается как статическая задача для каждого фиксированного среза ($y = t_k = \mathrm{const}$).

4. Целью исследований в работе является разработка оригинального нейросетевого статического логистического метода (НСЛМ) моделирования банкротств в сложных условиях зашумленности и неполноты данных с оптимальным отбором факторов для формирования исходных данных, а также агрегированием факторов в виде обобщенных функций желательности Харрингтона [11]. Цель направлена на повышение качества НСЛМ-моделей.

5. Для достижения цели было необходимо:

А. Решить задачу формализации и оптимизации отбора факторов в исходных данных.

Б. Изучить возможность эффективного использования в построении НСМ банкротств агрегирующих функций Харрингтона.

В. Разработать алгоритм регуляризации решения обратной задачи обучения НС, оценить в этом алгоритме влияние агрегирования факторов в виде обобщенных функций Харрингтона.

2. Результаты исследований. Количественные оценки

Сформулируем идеи, связанные с решением задач А, Б, В по разработке НСЛМ с оптимальным отбором факторов, а также с агрегированием факторов с указанием конкретной научной и прикладной новизны этих идей, а затем уже детализируем соответствующие операции алгоритмов указанных выше задач и приведем количественные оценки.

Задача А. Здесь осуществляется оптимальный отбор факторов. Формируется экспертным способом стартовый набор факторов, т. е. "сырые данные" D и по ним строится байесовский ансамбль вспомогательных нейросетевых субмоделей (ВНСМ), где НС-гипотезы $\{h_q(\mathbf{x}, \mathbf{W}, \mathbf{s})\}$ отличаются друг от друга видом активационных функций и параметрами структуры $\mathbf{s}$ (числом скрытых слоев нейронов и оптимальным числом нейронов в этих слоях). Согласно байесовской концепции регуляризации обучения [4, 10] все НС-гипотезы $\{h_q\}$, $q = \overline{1,Q}$ о происхождении данных D должны относиться к одному и тому же классу Ω, т. е. иметь одну и ту же "метагипотезу" $\forall h_q \in \Omega, \quad q = \overline{1,Q}$. В данной работе для выбора Ω мы использовали парадигму многослойных персептронов с обратным распространением ошибки при обучении (Multilayer Perseptron — Back Propagation (MLP-BP)). В алгоритме оптимального отбора факторов используется критерий $\bar{\Theta}$, усредненный на апостериорном, т. е. отфильтрованном байесовском ансамбле НС $\left\{h_q \middle| D, \mathbf{W}, H\right\}, \quad q = \overline{1,Q^*}, \quad Q^* \leqslant Q$. При этом критерий отбора $\bar{\Theta}$ выражается прямым способом через число ошибок первого и второго рода идентификации сетью объектов типа "банкрот — не банкрот". Организуется итерационный процесс оптимального отбора факторов, т. е. двойной фильтрации факторов исходных данных D: "внешних" итераций по индексу j (номеру фактора), $j = \overline{1,n}$ и "внутренних" итераций по индексу q (номеру сети) в байесовском ансамбле НС-гипотез $\{h_q\}$. Внешние итерации осуществляются как случайный выбор фактора из выборки D. Если фактор x_j оказывается достаточно информативным по критерию $\bar{\Theta}$ распознавания меток "банкрот — не банкрот" на выборке D, то он возвращается в выборку. В противном случае фактор исключается из выборки, поскольку он не несет полезной информации и ведет себя как шум. Такой случайный выбор охватывает перебор всех факторов. Критерий $\bar{\Theta}$ выражается следующими формулами:

$$\begin{aligned} &\Theta_{qj} = (N^{(\mathrm{I})}/N) + (N^{(\mathrm{II})}/N); \\ &\bar{\Theta}_j = \left[\sum_{q=1}^{Q^*} \Theta_{qj}\right] \Big/ N; \quad j = \overline{1,n}, \end{aligned} \tag{3}$$

где N — общее число корпораций-заемщиков в исходной выборке $\langle \mathbf{x}_j, y \rangle_{ig}$, $i = \overline{1,N}$, $g = \overline{1,G}$; $N^{(\mathrm{I})}$, $N^{(\mathrm{II})}$ — число ошибок первого и второго рода при идентификации обученной и протестированной НС предъявленных примеров по всей выборке; Q^* — число НСМ на байесовском ансамбле, отфильтрованном во внутреннем итерационном цикле по индексу $q = \overline{1,Q}$.

Правило отбора очередного j-го фактора в выборе:

$$x_j^* : \bar{\Theta}_j \leqslant \eta_1 \Big| \Theta_{jq} \leqslant \eta_2. \tag{4}$$

Здесь звездочка "*" в факторе x_j^* означает, что он возвращается в выборку D при отборе; η_1 и η_2 — экспертно задаваемые числа, характеризующие качество ансамблевой идентификации для НСМ. Если для банка-кредитора важно учесть в оценке (3)—(4) различие ошибок первого и второго рода с точки зрения кредитной политики, то можно это сделать введением коэффициентов предпочтения [3] Фишберна в формуле (3). Отметим, что критерий $\bar{\Theta}$, в отличие от критерия, принятого в [10], не требует априорного задания аналитического вида функции правдоподобия, что упрощает расчеты и приближает их к практике.

Во внутреннем итерационном процессе фильтрации НС-гипотез $\{h_q\}$, $q = \overline{1,Q}$ при каждом фиксированном факторе x_j, (j = const) используется тот же критерий отбора $\bar{\Theta}$ по (3)—(4). НС-гипотеза h_q остается в байесовском ансамбле, если для нее выполнено условие $\Theta_{jq} \leqslant \eta_2$. В противном случае h_q исключается для данного j-го шага внешних итераций из ансамбля.

Таким образом, оптимальный отбор факторов в итоге двойной фильтрации во внешнем (по индексу j) и внутреннем (по индексу q) цикле можно представить формулой

$$\mathbf{x}_{v.\text{опт}} = \arg\left[\min_{v=\overline{1,n}}\left\{\bar{\Theta}_v\right\} \middle| h_q(\mathbf{x}, y, \mathbf{W}) \in H\right], \tag{5}$$

где $\mathbf{x}_{v.\text{опт}}$ — оптимальный вектор (набор факторов) с номером v при случайном переборе всех факторов $(j = \overline{1,n})$; на этом наборе критерий $\bar{\Theta}_{\min}$ по (3)—(4) имеет минимальное значение.

Если множество $\mathbf{x}_{v.\text{опт}}$ не пусто, то (5) и есть мера эффекта предварительной компрессии факторов за счет их оптимального отбора. Поскольку алгоритм (5) включает в себя за счет внутренних итераций на байесовском ансамбле реализацию НСМ, то эту компрессию можно рассматривать как "предрегуляризацию" основной НСМ в итоге решения последующих задач Б и В.

Теперь мы можем на концептуальном уровне сформулировать научную новизну основной идеи, заложенной в алгоритм (3)—(5) решения задачи А, алгоритм отличается от известных алгоритмов следующими признаками.

1. Оптимальный отбор факторов осуществляется не изолированно от построения основной (рабочей) НСМ, а в тесной взаимосвязи с нею с использованием ансамбля ВНСМ из того же класса Ω и того же критерия качества (3), (5), что и в основной НСМ.

2. Одновременно с оптимальным отбором факторов осуществляется регуляризация ВНСМ на байесовском ансамбле НС, которую в составе НСЛМ можно трактовать как "предрегуляризацию" всей модели.

Порождаемый этими новыми признаками эмерджентный эффект — предварительная оптимальная компрессия факторного пространства и, соответственно, облегчение построения основной НСМ в аспекте обеспечения ее требуемых прогностических свойств.

Задача Б. Здесь осуществляется основной процесс комплексного (ступенчатого) агрегирования факторов, отличающийся от известных алгоритмов компрессии факторов тем, что вначале на эвристическом уровне по функциональному экономическому признаку формируются группы (кластеры) факторов, а затем внутри каждого u-го кластера факторы агрегируются в виде обобщенной функции желательности Харрингтона (ОФХ) [11] $\{H_u\}$; $u = \overline{1, M}$, и НС строится по агрегированным переменным в (2):

$$\hat{y}(H_u, t_k) = F(H_u, \mathbf{W}, t_k),\ k = \overline{1, N},\ u = \overline{1, M}. \quad (6)$$

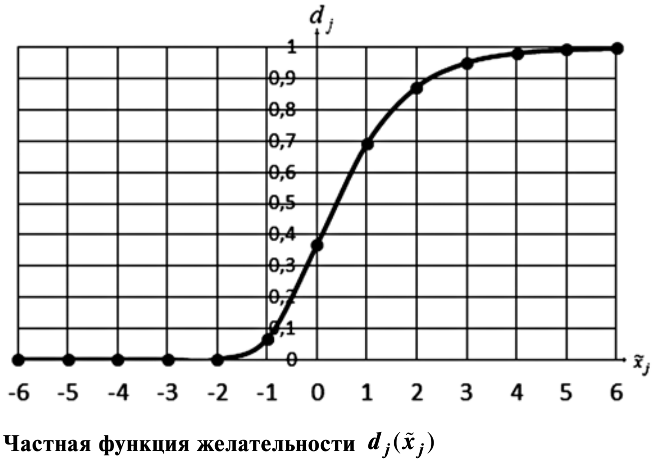
Достигаемые за счет этого признака НСЛМ эмерджентные эффекты:

- возможность внесения в алгоритм предобработки данных априорной эвристической информации, аккумулирующей профессиональный опыт аналитика, разрабатывающего НС при образовании кластеров факторов;
- сильная компрессия данных и облегчение обучения НС в аспекте повышения ее прогностической силы согласно рекомендациям Риссанена [6];
- возможность учета в явной форме нелинейного взаимного влияния факторов x_j, x_k, $k \neq j$ друг на друга внутри структур функции $\{H_u\}$;
- возможность наглядной экономической интерпретации зависимостей вероятности банкротства от агрегатов $\{H_u\}$, получаемых с помощью (1) и НСМ. Например, $\{H_u\}$ могут оценивать рентабельность бизнеса, ликвидность корпораций-заемщиков, платежеспособность и другие свойства.

Сделаем краткие комментарии по ОФХ. Более подробное описание содержится в работе [11]. Одним из наиболее удобных способов агрегирования как экзогенных x_j, так и эндогенных y_m, по мнению авторов статьи, является построение обобщенной функции желательности Харрингтона H. В основе построения этой функции лежит преобразование натуральных значений переменных в безразмерную шкалу желательности или предпочтительности: x_j, $y_m \to d_j$, где d_j — частная функция желательности.

Шкала желательности (см. рисунок) относится к психофизическим шкалам. Ее назначение — установление соответствия между физическими и психологическими (лингвистическими) величинами. Под психологическими величинами понимаются чисто субъективные (экспертные) оценки желательности того или иного значения переменных x_j, y_m. В данной работе все количественные оценки относятся к факторам $\{x_j\}$, $j = \overline{1, n}$, хотя полученные результаты в принципе справедливы и для агрегированных экзогенных переменных $\{y_m\}$.

Чтобы получить шкалу желательности, удобно пользоваться готовыми разработанными таблицами между отношениями предпочтения в числовой системе (отметка d_j на шкале частных функций желательности) и лингвистической шкале (табл. 1) [11].



Частная функция желательности $d_j(\tilde{x}_j)$

Таблица 1

Стандартные отметки на шкале желательности

Желательность (лингвистическая оценка переменной x_j)	Отметка на шкале желательности d_j
Очень хорошо	0,8...1,0
Хорошо	0,63...0,8
Удовлетворительно	0,37...0,63
Плохо	0,2...0,37
Очень плохо	0...0,2

Более точные результаты, чем табл. 1, дает аналитическое представление функции желательности:

$$d_j = \exp(-\exp(-\tilde{x}_j)); \; d_j \in [0,1]; \; \tilde{x}_j \in [-6,6]. \quad (7)$$

Заметим, что ширина интервала нормированной переменной $\tilde{x}_j$ должна выбираться так, чтобы значение $\tilde{x}_j = -6$ соответствовало в табл. 1 лингвистической оценке "Очень плохо" (d_j = 0), а $\tilde{x}_j = 6$ должно соответствовать оценке "Очень хорошо" (d_j = 1,0). Ось абсцисс на рисунке можно сузить, например $\tilde{x}_j \in [-3,3]$. Тогда функция желательности $d_j(\tilde{x}_j)$ будет иметь более крутой вид и точность графических оценок несколько снизится.

Харрингтоном был предложен единый комплексный мультипликативный показатель, а именно обобщенная функция желательности (среднее геометрическое в каждом u-м кластере):

$$H_u = \sqrt[n]{\prod_{j=1}^{n} d_j}, \; u = \overline{1, M}, \quad (8)$$

значения которой также могут быть интерпретированы по шкале, представленной в табл. 1. График ОФХ (8) будет таким же, как и для частных функций желательности d_j (7) (см. рисунок). В исследуемой нами задаче число агрегируемых факторов в каждом кластере будет различным, т. е. в (8) вместо n следует подставлять $\{n_u\}$.

Использование функции Харрингтона по (7), (8) требует выполнения следующих правил.

- Все факторы в фиксированном кластере факторов (u = const) должны быть однонаправленными, т. е. уменьшению вероятности банкротства P в логистической функции (1) должно соответствовать увеличение фактора x_j и, соответственно, рост показателя экспоненты $\hat{y}(\mathbf{x})$. Если некоторые из факторов $\{x_k\}$ в кластере имеют инверсный характер, т. е. рост этих факторов увеличивает вероятность банкротства P в формуле (1), то их следует преобразовать, например: $x_k \to x'_k = 1/x_k$.
- ОФХ H по (8) устроена так, что, если одна из частных функций желательности $d_j(\tilde{x}_j)$ внутри кластера $(x_j \in H_u, \; u = \text{const})$ равна нулю, то обнуляется и весь агрегат H_u с потерей его информативности. Поэтому авторы статьи, учитывая малую чувствительность функции частных желательностей $d_j(\tilde{x}_j)$ по (7) вблизи нуля, предлагают несколько загрубить оценку, принимая

$$d'_j = \begin{cases} \varepsilon, \text{ если } d_j \leqslant 0{,}1, \\ d_j \text{ в противном случае.} \end{cases} \quad (9)$$

Здесь ε — задаваемое аналитиком малое число, например ε = 0,05.

Задача B. По сути квазибайесовский алгоритм регуляризации (КБАР) решения ОЗ нахождения параметров НСМ, т. е. множества синаптических весов **W**, был описан выше в алгоритме задачи A. Более подробно этот вопрос изложен в работе [4]. Остается сформулировать основную идею КБАР и конкретизировать его новизну.

КБАР использует ту же парадигму регуляризации решения ОЗ, что и в теории А. Н. Тихонова — это сужение пространства искомых решений $Z' \subset Z$, где Z' — некоторый "компакт". Однако способ сужения Z до компакта Z' в КБАР другой, отличный от построения стабилизаторов Тихонова $\Omega(\mathbf{z})$.

В КБАР сужение пространства Z осуществляется последовательно в три этапа:

Этап 1. Выбор класса (метагипотезы Ω; $h_q(\mathbf{x}, y, \mathbf{W}) \in \Omega$) и вида НС-гипотез $\{h_q\}$ внутри Ω. Как было отмечено выше, в качестве Ω был выбран класс MLP-BP, а в качестве варьируемых внутри Ω активационных функций для $\{h_q\}$ — гиперболический тангенс или логистическая функция; в качестве параметров структуры $\boldsymbol{s}$ сети выбрано число скрытых слоев и оптимальное число нейронов в этих слоях.

Этап 2. Апостериорная фильтрация НС-гипотез байесовского ансамбля $\{h_q\}$, $q = \overline{1, Q}$ по прямому критерию ошибок первого и второго рода при идентификации объектов выборки в (4), (5).

Этап 3. Усреднение критерия Θ на отфильтрованном байесовском ансамбле Q^* по (3).

КБАР отличается от известного алгоритма С. А. Шумского [10] двумя новыми признаками:

1) отказом от допущения об известном априори виде закона распределения шумов (в [10] используются гауссовы и лапласовы распределения);

2) регуляризация решения ОЗ для основной НСМ выполняется взаимосвязано с алгоритмом оптимального отбора факторов: для одинакового класса $h_q \in \Omega$ вспомогательных субмоделей и основной модели, а также одинакового критерия качества (3), (4).

Получаемый эмерджентный эффект от этих новых признаков — упрощение расчетов (не требуется выполнять расчет функции правдоподобия распределения шумов, которая обычно неизвестна), обеспечение работоспособности (устойчивости) НСМ в сложных условиях моделирования (за счет процедур предрегуляризации — регуляризации ОЗ), приближение НСМ к практике.

Количественные оценки. В качестве исходных данных D использовали ретроспективные данные корпораций-заемщиков, одной из наиболее распространенных отраслей экономики — строительной отрасли, полученные фирмой "Bereua Van Dijk" [12]. База данных содержала 90 наблюдений. Использовалась система из 15 удельных показателей, широко применяемых в задачах оценки банкротств [13]: L_1 — быстрый коэффициент ликвидности; L_3 — коэффициент покрытия запасов; P_1 — текущий коэффициент ликвидности; F_1 — коэффициент финансовой зависимости; F_2 — коэффициент автономии собственных средств; F_3 — обеспеченность запасов собственными оборотными средствами; F_4 — индекс постоянного актива; R_1 — общая рентабельность; R_2 — рентабельность активов; R_3 — рентабельность собственного капитала; R_5 — рентабельность оборотных активов; A_2 — оборачиваемость активов; A_4 — оборачиваемость кредиторской задолженности; A_5 — оборачиваемость дебиторской задолженности; A_6 — оборачиваемость запасов.

Формулы для расчета этих удельных показателей содержатся в работе [3] и оперируют с данными стандартной бухгалтерской отчетности.

Был реализован алгоритм (5), описанный выше для задачи А оптимального отбора факторов с применением регуляризации на байесовском ансамбле ВНСМ.

Характеристики ансамбля: метагипотеза Ω — MLP-BP, реализованный на программном продукте NeuroSolutions 5.0 (демоверсия). Характеристики НС-гипотез $\{h_q\}$, $q = \overline{1,Q}$ приведены в

Таблица 2

Байесовский ансамбль НСМ $\{h_q\}$

№ НСМ	Число скрытых слоев	Вид активационной функции в скрытых слоях
1	1	(10)
2	2	(10) в обоих слоях
3	2	(10) в первом слое, (11) — во втором слое
4	1	(11)
5	2	(11) в обоих слоях
6	2	(11) в первом слое, (10) — во втором слое

табл. 2. В скрытых слоях НС использованы активационные функции логистической сигмоиды и гиперболического тангенса:

$$f(s) = \frac{1}{1 + e^{-as}}, \; a > 0; \tag{10}$$

$$f(s) = \operatorname{th}(bs), \quad b > 0. \tag{11}$$

Активационная функция выходного слоя во всех НСМ — линейная.

Указанные выше характеристики ансамбля были едиными в количественных оценках для всех трех задач А, Б, В (см. ниже).

Экспертно задаваемое число η_1 в правиле отбора факторов выбирали равным 0,1 (или 90 % правильно идентифицируемых объектов). Результаты оптимального отбора факторов показаны в табл. 3. Видно, что все шесть НС-гипотез $\{h_q\}$ в байесовском ансамбле успешно прошли фильтрацию и усредняемый критерий качества НСМ $\bar{\Theta}$ по (3) вычисляется по всем шести НС, т. е. $Q^* = 6$. Оптимальный вариант отбора соответствует пятой итерации и соответствует набору из 11 факторов: L_1, L_2, F_1, F_2, R_1, R_2, R_3, A_2, A_4, A_6 при минимальном значении критерия качества по (3) $\bar{\Theta}_{\min} = 0{,}0444$ (или 95,56 % верно идентифицированных объектов).

Таким образом, применение байесовского ансамбля в сочетании с отбором признаков приводит к улучшению модели: критерий $\bar{\Theta}$ уменьшился по сравнению с исходным значением ($\nu = 0$), равным 0,0778, до значения $\bar{\Theta} = 0{,}0444$, т.е. в 1,75 раз, при сокращении числа факторов с 15 до 11 (в 1,36 раз). Это весьма ощутимый выигрыш за счет оптимального отбора факторов, причем одновременно реализуется процесс предрегуляризации основной НСМ.

Таблица 3

Итерационный процесс оптимального отбора факторов

№ шага итерации v	Набор факторов $\{x_{vj}\}$	Исключение (−) или включение (+) факторов	$N^{(I)}$	$N^{(II)}$	$\bar{\Theta}$ по (3)
0	$L_1, L_2, P_1, F_1, F_2, F_3, F_4, R_1, R_2, R_3, R_5, A_2, A_4, A_5, A_6$	—	4	3	0,0778
1	$L_1, L_2, P_1, F_1, F_2, F_3, F_4, R_1, R_2, R_3, A_2, A_4\ A_5, A_6$	$-R_5$	2	3	0,0556
2	$L_1, L_2, P_1, F_1, F_2, F_3, R_1, R_2, R_3, A_2, A_4\ A_5, A_6$	$-F_4$	3	1	0,0444
3	$L_1, L_2, P_1, F_1, F_2, F_3, R_1, R_2, R_3, A_4\ A_5, A_6$	$-A_2$	4	1	0,0556
4	$L_1, L_2, P_1, F_1, F_2, F_3, R_1, R_2, R_3, A_2, A_4, A_6$	$+A_2, -A_5$	3	1	0,0444
5	$L_1, L_2, F_1, F_2, F_3, R_1, R_2, R_3, A_2, A_4, A_6$	$-P_1$	3	1	0,0444
6	$L_1, L_2, F_1, F_2, R_1, R_2, R_3, A_2, A_4, A_6$	$-F_3$	3	3	0,0667

Таблица 4

Результаты нейросетевого моделирования при агрегировании факторов на основе функций Харрингтона и сравнения с другими моделями

Номер модели	Показатель	Номер НС-гипотез в ансамбле						Среднее на ансамбле
		1	2	3	4	5	6	
1	$N^{(I)}$	2	7	5	0	2	2	1
	$N^{(II)}$	4	8	4	1	3	3	2
	Θ	0,0667	0,1667	0,1	0,0111	0,0556	0,0556	0,0333
2	$N^{(I)}$	4	4	3	3	2	3	4
	$N^{(II)}$	3	4	2	2	4	2	3
	Θ	0,0778	0,0889	0,0556	0,0556	0,0667	0,0556	0,0778
3	$N^{(I)}$	3	4	4	2	1	3	3
	$N^{(II)}$	2	3	1	1	2	1	1
	Θ	0,0556	0,0778	0,0556	0,0333	0,0333	0,0444	0,0444
4	$N^{(I)}$	2	3	2	2	1	1	2
	$N^{(II)}$	12	11	10	7	5	9	8
	Θ	0,1556	0,1556	0,1333	0,1	0,0667	0,1111	0,1111

Количественные оценки для задач Б и В сведены в итоговую табл. 4. Предлагаемый метод агрегирования на базе ОФХ (модель 1), который был подробно описан выше в задачах Б и В, сравнивался с широко применяемыми другими моделями: вариантами расчета на 15 факторах (без агрегирования, модель 2), с рассмотренным выше оптимальным отбором факторов (модель 3) и также с агрегированием факторов по методу нечеткой матричной свертки [13] (модель 4).

Заметим, что эффект фильтрации на апостериорном байесовском ансамбле (обученных и протестированных НС) проявился четко для модели 1 (отсеяна НС с номером 2 при $\Theta > \eta = 0{,}1$) и в модели 4 (НС с номерами 1, 2, 3 и 6).

Наилучший результат $\left(\bar{\Theta} = 0{,}0333\right)$ соответствует предлагаемой модели 1 с агрегированием факторов по четырем кластерам по ОФХ $\{H_u\}$, $u = 1, 2, 3, 4$. В кластере H_1 агрегируются факторы ликвидности; в кластере H_2 — факторы финансовой устойчивости; в кластере H_3 — факторы рентабельности и в кластере H_4 — факторы деловой активности. Таким образом, позиционированный эмерджентный эффект предлагаемого метода в обеих задачах Б и В подтвержден на реальных зашумленных данных строительной отрасли.

Заключение

В предложенном методе оптимального отбора факторов при формировании базы исходных данных D (см. табл. 3) достигается эммерджентный эффект — улучшение качества НСМ Θ по (5) в 1,75 раза при компрессии факторов в 1,35 раз, что достигается за счет взаимосвязи алгоритма отбора факторов с регуляризацией ВНСМ на байесовском ансамбле. Механизм указанной "взаимосвязи", по мнению авторов статьи, состоит в учете нелинейного влияния факторов друг на друга в процессе обучения ВНСМ в методе отбора факторов, что повышает информативность данных D и качество всей рабочей НСМ по прогностическому критерию $\bar{\Theta}$.

В предложенном методе комплексного (ступенчатого) агрегирования факторов на основе функций Харрингтона достаточно высокое прогностическое качество НСМ в сложных условиях моделирования банкротств (см. табл. 4) обеспечивается основным новым признаком метода: операция агрегирования проводится взаимосвязанно с операцией регуляризации решения ОЗ на байесовском ансамбле путем поэтапного сужения пространства искомых решений ($Z \to Z'$, где Z' — компакт).

Предложенные теоретические идеи двух оригинальных методов компрессии переменных подтверждены результатами на реальных данных строительной отрасли.

В аспекте дальнейшего развития предложенных двух методов представляет интерес их последовательное сочетание при большом числе факторов (сотни) — вначале оптимального отбора факторов, затем агрегирования факторов на основе функций Харрингтона.

Список литературы

1. **Галушкин А. И.** Нейронные сети: основы теории. М.: Горячая линия — Телеком, 2012. 469 с.
2. **Галушкин А. И.** Применение нейрокомпьютеров в финансовой деятельности. URL: http://www.hardline.ru/3/37/1484/ (дата обращения: 28.08.2017).
3. **Белолипцев И. И., Горбатков С. А., Романов А. Н., Фархиева С. А.** Моделирование управленческих решений в сфере экономики в условиях неопределенности. М.: ИНФРА-М, 2015. 299 с.
4. **Горбатков С. А., Полупанов Д. В., Макеева Е. Ю., Бирюков А. Н.** Методологические основы разработки нейросетевых моделей экономических объектов в условиях неопределенности. М.: Издательский дом "Экономическая газета", 2012. 494 с.
5. **Матвеев М. Г., Свиридов А. С., Алейникова Н. А.** Модели и методы искусственного интеллекта. Применение в экономике: учеб. пособие. М.: Финансы и статистика; ИНФРА-М, 2008. 448 с.
6. **Rissanen J.** Modeling by shortest data description // Automatica. 1978. Vol. 14. P. 465—471.
7. **Доленко С. А.** Нейросетевые методы решения обратных задач // XV Всероссийская научно-техническая конференция "Нейроинформатика — 2013": Лекции по нейроинформатике. М.: НИЯУ МИФИ, 2013. С. 214—269.
8. **Ohlson J. A.** Financial Ratios and the Probabilistic Prediction of Bankruptcy // Journal of Accounting Research, 1980, № 18 (1), p. 109—113.
9. **Тихонов А. Н., Арсенин В. Я.** Методы решения некорректных задач. М.: Наука. Физматлит, 1986. 288 с.
10. **Шумский С. А.** Байесова регуляризация обучения // Научная сессия МИФИ 2002. IV Научно-техническая конференция "Нейроинформатика — 2002": Лекции по нейроинформатике. Часть 2. М.: МИФИ, 2002. С. 30—93.
11. **Адлер Ю. П., Маркова Е. В., Грановский Ю. В.** Планирование эксперимента при поиске оптимальных условий. М.: Наука, 1976. 279 с.
12. **Makeeva E. U., Neretina E. A.** Binary model versus discriminant analysis relating to corporate bankruptaes: The Case of Russian Construction Industry // Journal of Accounting, Finance and Economics. 2013. Vol. 3. № 1. P. 65—76.
13. **Шевченко И. В., Халафян А. А., Васильева Е. Ю.** Создание виртуальной клиентской базы для анализа кредитоспособности российских предприятий // Финансы и кредит. 2010. № 1 (385). С. 13—18.

S. A. Gorbatkov, D. Sc. Professor, sgorbatkov@mail.ru,
Financial university under the government of the Russian Federation, Ufa (branch), Ufa 450015, Russia,
D. V. Polupanov, Ph. D., Associate Professor, demetrious@mail.ru,
Bashkir State University, Ufa, 450076, Russia

Optimal Selection and Aggregation of Exogenous Variables in Neural Network Models of Bankruptcies Based on Harrington Functions

The article is devoted to the issue of constructing and further improving neural network models in conditions of noise, inaccuracy, incompleteness and uncertainty of the data. A practical application is the development of an original neural network logistics method for assessing the probability of enterprise bankruptcies. In order to improve the quality, adequacy, accuracy and predictive properties of neural network models, two methods of reducing factor space are proposed: 1) an iterative method of optimal factor selection; 2) aggregation of factors using the generalized Harrington desirability function. A common original concept of the proposed methods, which differs from the known methods of exogenous variables compression, is a complex (systemic) study of the interrelation between the operations of optimal factor selection and their aggregation with *the operation of regularizing the inverse problem of training neural networks in a Bayesian ensemble. The generated emergent effect is the compression of the factor space and, accordingly, facilitating the construction of a neural network model in terms of providing its required prognostic properties. The proposed theoretical ideas of the two original methods of variable compression are confirmed by the results on real data of construction industry enterprises. Quantitative estimates are presented comparing the proposed methods with the construction of neural network models on the initial set of factors, as well as with the aggregation of factors by the fuzzy matrix convolution method.*

Keywords: *optimal selection, aggregation of factors, neural network models of bankruptcies, system approach to modeling*

References

1. **Galushkin A. I.** *Nejronnye seti: osnovy teorii* [Neural networks: the fundamentals of the theory], Moscow, Gorjachaja linija — Telekom, 2012. 469 p. (in Russian).

2. **Galushkin A. I.** *Primenenie nejrokomp'juterov v finansovoj dejatel'nosti* [The use of neurocomputers in financial activities]. URL: http://www.hardline.ru/3/37/1484/ (data of access: 28.08.2017). (in Russian).

3. **Belolipcev I. I., Gorbatkov S. A., Romanov A. N., Farhieva S. A.** *Modelirovanie upravlencheskih reshenij v sfere jekonomiki v uslovijah neopredelennosti*, Moscow, INFRA-M, 2015, 299 p. (in Russian).

4. **Gorbatkov S. A., Polupanov D. V., Makeeva E. U., Birjukov A. N.** *Metodologicheskie osnovy razrabotki nejrosetevyh modelej jekonomicheskih objektov v uslovijah neopredelennosti* [The methodological basis for the development of neural network models of economic entities in conditions of uncertainty], Moscow, Izdatel'skij dom "Ekonomicheskaja gazeta", 2012, 494 p. (in Russian).

5. **Matveev M. G., Sviridov A. S., Alejnikova N. A.** *Modeli i metody iskusstvennogo intellekta. Primenenie v jekonomike: uchebnoe posobie.* [Models and methods of artificial intelligence. Application in Economics: Textbook], Moscow, Finansy i statistika; INFRA-M, 2008, 448 p. (in Russian).

6. **Rissanen J.** Modeling by shortest data description, *Automatica*, 1978, vol. 14, pp. 465—471.

7. **Dolenko S. A.** Nejrosetevye metody reshenija obratnyh zadach [Neural network methods for solving inverse problems], *XV Vserossijskaja nauchno-tehnicheskaja konferencija "Nejroinformatika — 2013": Lekcii po nejroinformatike*, Moscow, NIJaU MIFI, 2013, pp. 214—269 (in Russian).

8. **Ohlson J. A.** Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, 1980, no. 18 (1), pp. 109—113.

9. **Tihonov A. N., Arsenin V. Ja.** *Metody reshenija nekorrektnyh zadach* [Methods for solving ill-posed problems]. Moscow, Nauka. Fizmalit, 1986, 288 p. (in Russian).

10. **Shumskij S. A.** Bajesova reguljarizacija obuchenija [Bayesian regularization of learning], *Nauchnaja sessija MIFI 2002. IV Nauchno-tehnicheskaja konferencija "Nejroinformatika — 2002": Lekcii po nejroinformatike. Chast' 2*, Moscow, MIFI, 2002, pp. 30—93 (in Russian).

11. **Adler Ju. P., Markova E. V., Granovskij Ju. V.** Planirovanie jeksperimenta pri poiske optimal'nyh uslovij. [Planning an experiment in searching for optimal conditions], Moscow, Nauka, 1976, 279 p. (in Russian).

12. **Makeeva E. U., Neretina E. A.** Binary model versus discriminant analysis relating to corporate bankruptaes: The Case of Russian Construction Industry, *Journal of Accounting, Finance and Economics*, 2013, vol. 3, no. 1, pp. 65—76.

13. **Shevchenko I. V., Halafjan A. A., Vasil'eva E. Ju.** Sozdanie virtual'noj klientskoj bazy dlja analiza kreditosposobnosti rossijskih predprijatij [Creating a virtual customer base for the analysis of the creditworthiness of Russian companies], *Finansy i kredit*, 2010, no. 1 (385), pp. 13—18 (in Russian).