

С. А. Горбатков, С. А. Фархиева, М. В. Коротнева, А. А. Чапкович (Уфа, филиал ВЗФЭИ, УГАТУ). Метод вложенных математических моделей для регуляризации нейросетевых задач финансового контроля.

Эффективным инструментом решения некорректных по Адамару задач восстановления при помощи нейросетевых отображений является метод регуляризации по Тихонову [1]. В докладе обсуждается возникающая при этом проблема состоятельности регуляризованной обратной задачи. Обозначим Z метрическое пространство искомых характеристик объекта (процесса), а U — пространство косвенных характеристик объекта. Пусть каждому $z \in Z$ при помощи оператора A ставятся в соответствие элементы $u \in U$: $Az = u$, где A — оператор нейросетевого отображения.

Пусть Z^* — компакт [1] и $z \in Z^*$. Тогда естественной является постановка задачи восстановления на компакте Z^* :

$$z_\delta: \rho_{\tilde{U}}(Az, \tilde{u}) \leq \delta, \quad z \in Z^*, \quad z_\delta \in Z_\delta, \quad (1)$$

где \tilde{u} — возмущенный элемент характеристик объекта; δ — мера погрешности данных; \tilde{u} , δ известны. Задача (1) названа в [1] задачей квазиминимизации $\rho_{\tilde{U}}$. Утверждается, что всякое решение z_δ функционального неравенства (1) удовлетворяет принципу регуляризации Тихонова [1] и является устойчивым приближением к искомому решению z . При этом может возникнуть несостоятельность задачи (1): $\rho_0 = \inf \rho_{\tilde{U}}(Az, \tilde{U}) > \delta$. Нами предлагается метод вложенных математических моделей (ВММ), позволяющий за счет специальной предобработки данных уменьшать меры δ и ρ_0 взаимосвязано так, что функциональное неравенство (1) будет разрешимым. Ядром метода ВММ является итерационный процесс оптимальной отбраковки аномальных наблюдений в данных:

$$i^{*(k+1)}: \left[\max_{i \in N^{(k)}} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| 100\% \right] > \eta_i, \quad i \in \Omega_{learn} \cup \Omega_{test}, \quad k = 1, 2, \dots,$$

правило остановки итераций: $\delta(N^{(k)}, E^{(k)}) \rightarrow \min_k \delta$, $N^{(k)} \geq \xi n$, $\delta = C_1 E + C_2 (NMSE)$, где $i^{*(k+1)}$ — номер точки, отбраковываемой на $(k+1)$ -итерации; $N^{(k)}$ — количество точек на k -й итерации в обучающем множестве; η_i — уровень относительной ошибки; ξ — экспериментально подбираемый коэффициент репрезентативности выборки ($\xi \approx 5 \div 10$); n — число входных факторов в модели; C_1, C_2 — весовые множители, задаваемые экспертно ($C_1 = C_2 = 0,5$); $NMSE$ — нормированное среднее квадратическое значение остатков в расчете; E — ошибка обобщения.

Для вычисления δ на каждом шаге итераций строится вспомогательная нейросетевая субмодель (ВММ). В отличие от известных методов отбраковки данных, в описанном выше методе меры $\delta^{(k)}$ и $\rho_0^{(k)}$ связаны через алгоритм модификации синаптических весов в сети [2].

Метод ВММ апробирован в [2] на реальных данных налогового контроля.

СПИСОК ЛИТЕРАТУРЫ

1. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. М.: Наука, Физмалит, 1979.
2. Горбатков С. А., Полупанов Д. В. Методы нейроматематики в налоговом контроле. Уфа: РИЦ БашГУ, 2008.