Объединяя полученные двухклассовые классификаторы в один, получаем трехклассовый классификатор. Для этого воспользуемся тем, что классы «здоровая растительность» c_h и «пораженная растительность» c_d являются подклассами класса «растительность» c_p. Значение на выходе классификатора «растительность–почва», соответствующее классу «растительность» , находится в диапазоне [0; 1] и может быть использовано в качестве коэффициента для выходов классификатора «здоровая–пораженная растительность».

Тестирование объединения классификаторов осуществлялось на общей базе изображений классов (обучающая и валидационная выборка), которое показало снижение влияния шумовых факторов (освещение, солнечные блики) на качество получаемых карт пораженной растительности.

Полученные результаты тестирования приведены в таблице:

Класс	Пораженная растительность	Здоровая растительность	Почва
Точность, %	75.61	85.65	99.03

Наибольшее количество ошибок возникало на участках, соответствующих границе здоровой растительности и почвы (в особенности в местах, где небольшие участки почвы окружены растительностью, создающей на этом участке почвы тень). При этом использование двух СНС позволило снизить искажение цветовых характеристик изображений окрестностей пикселей при попадании в него растительности и почвы.

Список цитированных источников

1. Николенко, С. Глубокое обучение / С. Николенко, А. Кадурин, Е. Архангельская. – СПб.: Питер, 2018. – 480 с.

Работа выполнена при частичной финансовай поддержке БРФФИ (проект № Ф18В-005) и ГКНТ Республики Беларусь (проект № Ф18ПЛШГ-008П).

УДК 004.021:032.26

О СКОРОСТИ СХОДИМОСТИ И ВЫБОРЕ ШАГА ОБУЧЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ПРЯМОГО РАСПРОСТРАНЕНИЯ

Гладкий И.И., Будяков В.В., Чикалов Б.И., Юхимук Т.Ю.

Брестский государственный технический университет, г. Брест

Научный руководитель: Махнист Л.П., канд. техн. наук, доцент

Задача обучения нейронной сети прямого распространения состоит в нахождении весовых коэффициентов w_{ij} и порогов T_j нейронной сети, которые минимизируют функцию ошибки сети

$$E\left(w_{11}, w_{21}, \ldots, w_{m1}, T_1, \ldots, w_{1n}, w_{2n}, \ldots, w_{mn}, T_n\right) = \frac{1}{2}\sum_{j=1}^{n}\left(y_j - t_j\right)^2,$$

где $y_j = F\left(S_j\right)$ – значение функции активации j-ого выходного нейрона сети, $S_j = \sum_{i=1}^{m} w_{ij}x_i - T_j$, x_i – выходное значение i-ого нейрона предыдущего слоя, t_j – ожидаемый выход j-ого выходного нейрона ($i = \overline{1,m}$, $j = \overline{1,n}$).

Введем обозначения: $\overline{W} = \left(w_{11}, w_{21}, K, w_{m1}, T_1, K, w_{1n}, w_{2n}, K, w_{mn}, T_n\right)^T$ – вектор-столбец весовых коэффициентов w_{ij} и порогов T_j нейронной сети, а $\overline{W}_j = \left(w_{1j}, w_{2j}, K, w_{mj}, T_j\right)^T$ – вектор-столбец весовых коэффициентов w_{ij} и порога T_j, связанных с j-ым выходным нейроном сети, $E\left(\overline{W}\right) = \frac{1}{2}\sum_{j=1}^{n}\left(y_j - t_j\right)^2$ – функция ошибки сети, $E\left(\overline{W}_j\right) = \frac{1}{2}\left(y_j - t_j\right)^2$ – функция ошибки j-ого выходного нейрона сети.

Обучение нейронной сети с использованием метода наискорейшего спуска состоит в изменении весовых коэффициентов w_{ij} и порогов T_j нейронной сети на каждом $(t+1)$ шаге обучения ($t = 1, 2, K$) в соответствии с формулами:

$$\overline{W}_j(t+1) = \overline{W}_j(t) - a_j(t)\nabla E\left(\overline{W}_j(t)\right),$$

если шаг обучения $a_j(t)$ выбирается только для минимизации функции ошибки сети $E\left(\overline{W}_j\right)$, где $\nabla E\left(\overline{W}_j\right)$ – градиент функции $E\left(\overline{W}_j\right)$,

$$\overline{W}_j(t+1) = \overline{W}_j(t) - a(t)\nabla E\left(\overline{W}_j(t)\right),$$

если шаг обучения $a(t)$ выбирается для минимизации функции ошибки сети $E\left(\overline{W}\right)$.

Для выбора шага обучения предлагается использовать соотношения:

$$a_j(t) = \frac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)}, \tag{1}$$

$$a(t) = \frac{\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)}, \tag{2}$$

где $\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|$ – длина вектора градиента $\nabla E\left(\overline{W}_j(t)\right)$, связанная со скалярным произведением $\left(\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)$, а $\nabla^2 E\left(\overline{W}_j(t)\right)$ – матрица Гессе функции $E\left(\overline{W}_j(t)\right)$.

Соотношения (1) и (2) получены и использовались, например, в [1–3] при выполнении условия $\min_j\left(F'\left(S_j(t)\right)\right)^2 + \left(y_j(t) - t_j\right)F''\left(S_j(t)\right) > 0$, которое рассматривалось в [4].

Заметим, что выражения (1) и (2) можно записать в виде:

$$a_j(t)=\frac{1}{\left(\sum_{i=1}^{m}x_i^2+1\right)\cdot\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)},$$

$$a(t)=\frac{\sum_{j=1}^{n}\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2}{\left(\sum_{i=1}^{m}x_i^2+1\right)\cdot\sum_{j=1}^{n}\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)}.$$

Установлено, что величины $a_1(t)$, $a_2(t)$,..., $a_n(t)$ и $a(t)$, связаны соотношением

$$a(t)=\frac{1}{\sum_{j=1}^{n}\frac{q_j(t)}{a_j(t)}},\ \text{где}\ q_j(t)=\frac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2}.$$

Следовательно, $a(t)$ – взвешенное среднее гармоническое величин $a_j(t)$, с весами $q_j(t)$.

Полученное соотношение можно пояснить, например, следующим образом: если путь $\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2=\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2$ разбит на участки $\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2$, скорость (шаг обучения) на каждом из которых постоянна и равна $a_j(t)$, то средняя скорость (шаг обучения) $a(t)$ по всему пути $\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2$ будет равна взвешенному среднему гармоническому скоростей (шагов обучения) $a_j(t)$.

Сравним абсолютные изменения функции ошибки сети $E\left(\overline{W}\right)$ в рассмотренных двух случаях.

В первом случае получим, что абсолютное изменение функции ошибки сети $E\left(\overline{W}\right)$ удовлетворяет соотношению:

$$\left|\Delta_1 E\left(\overline{W}(t)\right)\right|=\left|E\left(\overline{W}(t+1)\right)-E\left(\overline{W}(t)\right)\right|=\left|\sum_{j=1}^{n}\left(E\left(\overline{W}_j(t+1)\right)-E\left(\overline{W}_j(t)\right)\right)\right|=$$

$$=\frac{1}{2}\sum_{j=1}^{n}a_j(t)\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2=\frac{1}{2}\left(\sum_{j=1}^{n}q_j(t)a_j(t)\right)\cdot\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2,$$

а во втором случае – соотношению:

$$\left|\Delta_2 E\left(\overline{W}(t)\right)\right|=\left|E\left(\overline{W}(t+1)\right)-E\left(\overline{W}(t)\right)\right|=\frac{a(t)}{2}\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2=\frac{1}{2\sum_{j=1}^{n}\frac{q_j(t)}{a_j(t)}}\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2.$$

Используя неравенство Йенсена, получим:

$$\max\left(\left|\Delta_1 E\left(\overline{W}(t)\right)\right|, \left|\Delta_2 E\left(\overline{W}(t)\right)\right|\right) = \frac{1}{2}\max\left(\sum_{j=1}^{n} q_j(t)a_j(t), \frac{1}{\sum_{j=1}^{n}\frac{q_j(t)}{a_j(t)}}\right)\cdot\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2 =$$

$$= \frac{1}{2}\left(\sum_{j=1}^{n} q_j(t)a_j(t)\right)\cdot\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2.$$

Следовательно, использование соотношений (1) будет обеспечивать лучшую скорость сходимости обучения нейронной сети с использованием метода наискорейшего спуска по сравнению с использованием соотношения (2).

Список цитированных источников

1. Golovko, M. Multilayer neural networks training methodic / M. Golovko, L. Makhnist, N. Maniakov // Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2003) : Proceedings, Lviv, Ukraine, 8–10 Sept. 2003. – Lviv, 2003. – P. 185–190.
2. Makhnist, L. Some Methods of Adaptive Multilayer Neural Network Training / L. Makhnist, N. Maniakov // International Journal of Computing. – 2004. – Vol. 3. – P. 99–106.
3. Maniakov, N. Traing algorithm for forecasting multilayer neural network / N. Maniakov, L. Makhnist, V. Rubanov // Pattern Recognition and Information Processing : Proceedings of The Seventh International Conferences (PRIP'2003), Minsk, Republic of Belarus, 21–23 May 2003 : in 2 vol. – Minsk, 2003. – Vol. 1. –P. 26–30.
4. Maxnist, L. Convergence Analysis of Neural Networks Training Based on Steepest Descent Method / L. Maxnist, A. Doudkin, V. Golovko // Pattern Recognition and Information Processing (PRIP'2007) : Proceedings of the Ninth International Conference, Minsk, Republic of Belarus, 22–24 May 2007 : in 2 vol. – Minsk, 2007. – Vol. 1. – P. 285–289.

УДК 004.832.2

РЕШЕНИЕ ЗАДАЧ ОПТИМИЗАЦИИ С ПРИМЕНЕНИЕМ ЭВОЛЮЦИОННОГО ПРОГРАММИРОВАНИЯ

Давидюк Ю. И.

Брестский государственный технический университет, г. Брест, Беларусь
Научный руководитель: Кочурко П. А., канд. техн. наук, доцент

В последнее время для решения каких-либо задач все чаще используют методы машинного обучения. Так, например, для решения оптимизационных задач пробуют использовать генетические алгоритмы, для задач классификации – нейронные сети. Для решения методами машинного обучения необходимо проанализировать задачу и найти наиболее эффективный алгоритм решения и способ настройки параметров для выбранного алгоритма.

Генетический алгоритм – это своеобразный метод оптимизации, основанный на идее естественного отбора как средства достижения наилучшего результата. Многочисленные методы естественной эволюции дают возможность решать разнообразные задачи.