

Автоассоциативные нейронные сети в задаче классификации с усеченным множеством

А. А. Хусаенов¹

Рассматривается модель оценки риска неблагоприятного клинического исхода². Предлагается применение метода обучения без учителя для задачи бинарной классификации с обучающим множеством, имеющим единственный ответ. Под усеченным множеством понимается не только малое количество примеров одного из случаев (благоприятного или неблагоприятного), но и данные полученные после очистки исходной таблицы. Некоторые результаты применения данной модели представлены в совместном исследовании [1], выполненном сотрудниками Национального медицинского исследовательского центра терапии и профилактической медицины МЗ РФ и механико-математического факультета МГУ имени М.В. Ломоносова.

Ключевые слова: нейронные сети, обучение без учителя, автоассоциативные нейронные сети, неблагоприятный клинический исход.

1. Введение

В исследовании [1] была сформулирована задача бинарной классификации клинических событий на основе анализов пациента. Рассматриваются события 2-х типов: благоприятное и неблагоприятное. Данные были представлены таблицей, содержащей 5062 строки и 66 столбцов.

Основной проблемой являлось малое количество данных неблагоприятных событий. Из 5062 доступных событий:

- 84 неблагоприятных
- 4978 благоприятных

¹Хусаенов Артем Азатович — аспирант кафедры математической теории интеллектуальных систем мех.-мат. ф-та МГУ; e-mail: a.khusaenov@mail.ru

Khusaenov Artem Azatovich — postgraduate student, Moscow State University, faculty of Mechanics and Mathematics, Mathematical Theory of Intelligent Systems department

²работа выполнена при поддержке РФФИ грант № 19-29-01051 «Разработка алгоритмов принятия решений для управления рисками неблагоприятных клинических событий в высокотехнологичной медицинской организации на основе технологии data mining»

Так как неблагоприятных событий мало, то при обучении модели на данном множестве возникает высокая вероятность ошибки первого рода (больной пациент признан здоровым). Возможно обучить модель со смещением наибольшей ошибки в зону второго рода (здоровый пациент признан больным). При этом ошибка останется допустимой. В таком случае допустимая доля здоровых пациентов будет лишней раз показана врачу.

Модель обучается на благоприятных событиях. При обучении выделяется главная компонента множества, а разделяющая поверхность сдвигается к границам этого множества. В случае возникновения отклонения модели выше некоторого порогового значения, событие признается неблагоприятным.

2. Постановка задачи и модель

Набор признаков, методы предобработки данных и устранения пропусков подробно описаны в исследовании [1]. Обозначим:

- $Y = (y_1, y_2, \dots, y_m)$ - множество ответов о событиях
- $X_i^n = (x_{1_i}, x_{2_i}, \dots, x_{n_i})$ - множество из n признаков события i
- $Y^0 = \{y_i = 0 \mid y_i \in Y, i = \overline{1, m}\}$ - класс благоприятных событий
- $Y^1 = \{y_i = 1 \mid y_i \in Y, i = \overline{1, m}\}$ - класс неблагоприятных событий

Объем данных после устранения и обработки пропусков:

- $n = 23$
- $|Y^0| = 2893$
- $|Y^1| = 66$

Модель

Применялась автоассоциативная нейронная сеть [2] (рис. 1):

- 23 нейрона входного слоя ($n = 23$)
- 12 нейронов скрытого слоя
- 23 нейрона выходного слоя ($n = 23$)

В качестве ответа (учителя) используется $X'_i = (x'_{1_i}, x'_{2_i}, \dots, x'_{23_i})$ - дубликат множества признаков X_i . То есть вместо решения задачи классификации $F : X_i \rightarrow y_i$ имеем $F : X_i \rightarrow X'_i$.

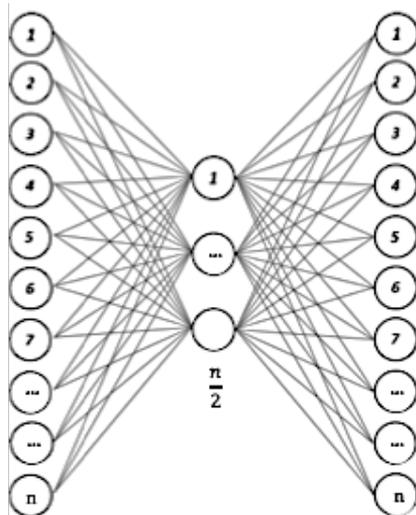


Рис. 1. архитектура нейронной сети

Нейронная сеть обучалась на множестве Y^0 и минимальное отклонение обучения по каждому выходному нейрону принималось как пороговое значение классификации. После обучения сети некоторое тестовое событие y_k оценивалось как:

- $y_k = 1$, если возникло отклонение выше порогового в одном из выходных нейронов
- $y_k = 0$, если отклонения по всем выходным нейронам не превышало порогового

Функции активации - сигмоидальные, функция потерь - разность по каждому выходному нейрону. Обучение производилось методом обратного распространения ошибки.

Результаты

Тестовое множество состояло из 132 событий (все события Y^1 и, измеримо, 66 событий из Y^0). Обучающее множество состояло из 2827 событий (все остальные события Y^0). На тестовом множестве достигнуто 100% точности.

При добавлении 12 несущественных признаков [1] достигнута точность 97%, где 3% - ошибка второго рода (4 здоровых пациента признаны больными).

Необходимо учитывать, что результаты получены на относительно небольшом тестовом множестве. Для объективной проверки необходим доступ к большему объему неблагоприятных событий и различные этапы повторных тестирований.

3. Смещение рода ошибки

Пусть A - множество нейронов скрытого слоя. Легко заметить, что в указанной модели (рис.2), для $F_1 : X_i \rightarrow A$ и $F_2 : A \rightarrow X'_i$ выполняется

$$F_2(A) = F_1^{-1}(X) + E^n, \quad (1)$$

где $E^n = (e_1, e_2, \dots, e_n)$ - ошибка обучения

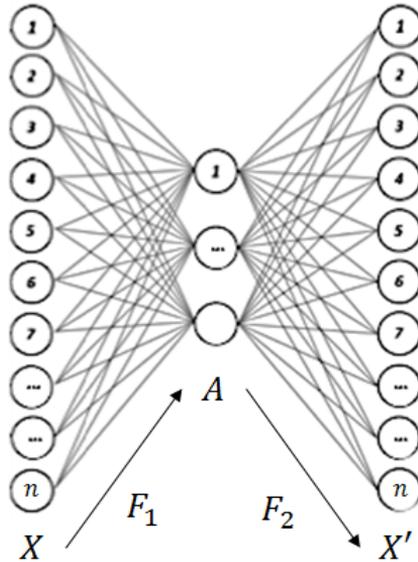


Рис. 2. Отображения в автоассоциативной нейронной сети

Будем считать, что нейронная сеть произвела обучение, если $E^n = \sigma^n$, где $\sigma^n = (\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_n)$ - неустранимая ошибка обучения [3].

Скрытый слой A значительно меньшей размерности (рис.2) усиливает значения, позволяющие осуществить обратное отображение $F_2(A)$ с наименьшей ошибкой $E^n \rightarrow \sigma^n$, и ослабляет значения вносящие шум. Тогда $F_2(A)$ при $E^n = \sigma^n$ будем называть *главной компонентой* множества X^n (рис. 3).

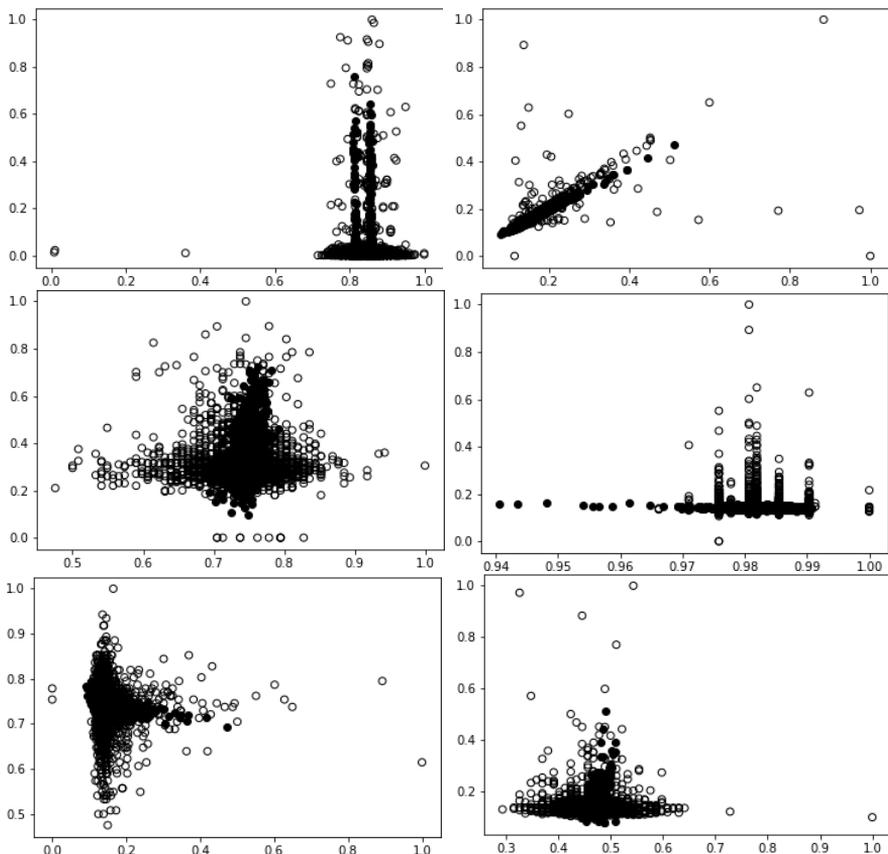
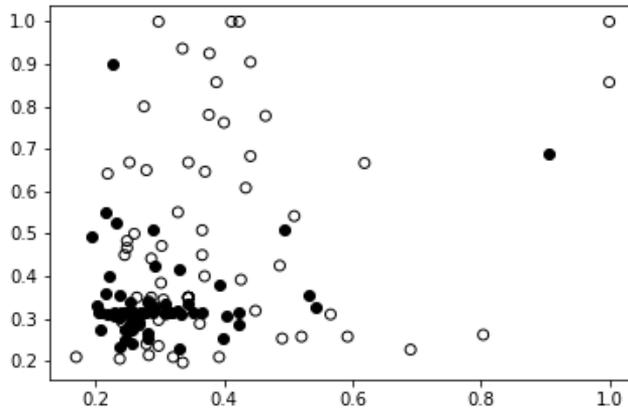


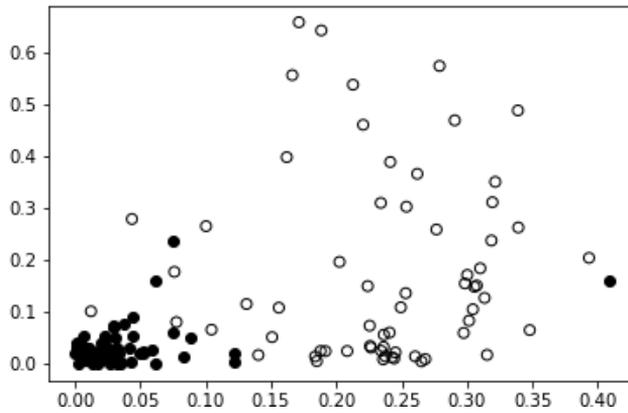
Рис. 3. главная компонента множества X^n в 2-мерных срезах
(черные точки - $F_2(A)$, белые точки - X)

Классификация тестового события y_k возможна в пространстве ошибок за счет оценки отклонений от главной компоненты $F_2(A)$. Поскольку обучение производилось на множестве Y^0 , то предполагается, что отклонение для событий Y^1 будет выше.

В пространстве ошибок события Y^0 будут сконцентрированы в зоне нуля, а Y^1 распределены по всему пространству. Произвести классификацию на пространстве ошибок будет проще, чем на пространстве признаков (рис. 4)



(а)



(б)

Рис. 4. (а) - пространство признаков в одном из 2-мерных срезов
 (б) - пространство ошибок для среза (а)
 (черные точки - Y^0 , белые точки - Y^1)

Классификация в пространстве ошибок (рис.4(б)) может быть реализована отдельными методами. В данной задаче было достаточно порогового условия (рис.5) вида

$$y_k = \begin{cases} 1, & \exists e_j > \sigma_j, e_j \in E_k \\ 0, & \forall e_j < \sigma_j, e_j \in E_k \end{cases} \quad (2)$$

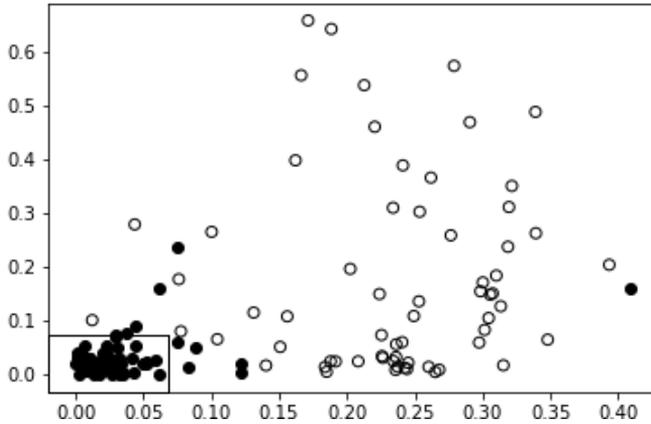


Рис. 5. Классификация в пространстве ошибок
(черные точки - Y^0 , белые точки - Y^1)

Ошибки второго рода здесь - это выбросы класса Y^0 . Уменьшение количества выбросов возможно за счет увеличения данных класса Y^0 (главная компонента множества будет сформирована более точно). При этом увеличить объем данных по здоровым пациентам проще, чем по больным.

Необходимо заметить, что на рис. 4 и рис. 5 представлены 2-мерные срезы. При классификации в многомерном виде ошибка значительно ниже (п.2 - результаты)

4. Отдельные замечания

Замечание 1. При обучении нейронной сети вида $F : X_i \rightarrow y_i$ на подобном множестве (п.1) риск переобучения крайне велик. Так как данных Y^1 значительно меньше Y^0 , то разделяющие поверхности могут распределиться как показано на рис.6(б).

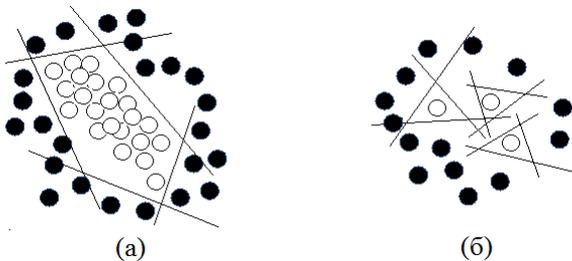


Рис. 6. (а) - допустимая обобщающая способность (б) - переобучение

Замечание 2. Данные должны быть нормализованы к области допустимых значений функций активации нейронов выходного слоя.

Замечание 3. Количество нейронов скрытого слоя A в данном случае подобрано эмпирически.

Выражаю благодарность проф. Рыжову А.П. и доц. Строгалову А.С. за обсуждения работы и ценные замечания.

Список литературы

- [1] Горный Б.Э., Рыжов А.П., Строгалов А.С., Журавлев А.Д., Хусаенов А.А., Шергин И.А., Фещенко Д.А., Абдуллаев А.М., Концевая А.В., “Оценка риска неблагоприятного клинического исхода методами углубленного анализа данных”, *Интеллектуальные системы. Теория и приложения*, **2:25** (2021).
- [2] Hinton G.E. , Salakhutdinov R.R., “Reducing the Dimensionality of Data with Neural Networks”, *Science*, **313** (2006), 504–507.
- [3] Stuart Geman, Bienenstock E., Doursat R., “Neural networks and the bias/variance dilemma”, *Neural Computation*, **1:4** (1992), 1–58.

Autoassociative neural networks in a classification problem with truncated dataset Khusaenov A.A.

The adverse clinical outcome risk assessing model is considering. It is proposed the unsupervised learning method application for a binary classification problem with a single answer training set. A truncated set is a dataset with a small examples number of one of the classes (favorable or unfavorable). A truncated set is also the data obtained after original table clearing. Some results of this model application are presented in a research [1] conducted by the scientists of the National Research Center for Therapy and Preventive Medicine of the Ministry of Health of the Russian Federation and the scientists of the Faculty of Mechanics and Mathematics of the Lomonosov Moscow State University. It is proposed the general method for such problems.

Keywords: neural networks, unsupervised learning, Auto-associative neural networks, autoencoder, adverse clinical outcome.

References

- [1] Gornyi B.E., Ryjov A.P., Strogalov A.S., Zhuravlev A.D., Khusaenov A.A., Shergin I.A., Feshchenko D.A., Abdullaev A.M., Kontsevaya A.V., “The adverse clinical outcome risk assessment by in-depth data analysis methods”, *Intelligent systems. Theory and applications*, **2:25** (2021) (In Russian).

- [2] Hinton G.E. , Salakhutdinov R.R., “Reducing the Dimensionality of Data with Neural Networks”, *Science*, **313** (2006), 504–507.
- [3] Stuart Geman, Bienenstock E., Doursat R., “Neural networks and the bias/variance dilemma”, *Neural Computation*, **1:4** (1992), 1–58.