

ФУНКЦИЯ КЛЕТОЧНОГО АНСАМБЛЯ – «ЕСТЕСТВЕННАЯ» КЛАССИФИКАЦИЯ ДЕЙСТВИТЕЛЬНОСТИ РЕЗОНАНСОМ ЗАКОНОМЕРНЫХ СВЯЗЕЙ

Е. Е. Витяев

Институт математики им. С.Л.Соболева СО РАН, vityaev@math.nsc.ru

Annotation. The notion of “natural” classification is introduced in the paper. Properties of “natural” classification as a law of nature are analyzed. The analogy between the “natural” classification and neurons ensemble functions demonstrated on the example of perception. Finally, the hypothesis about the function of neurons ensemble as a “natural” classification of reality by the resonance of regularities is proposed.

Введение

Понятие «естественной» классификации развивалось в СССР и России в рамках классификационного движения (аналога этому направлению нет в зарубежных исследованиях). В рамках этого направления был систематизирован опыт естествоиспытателей по созданию естественных классификаций и систематизированы критерии “естественности” классификации [6]. Приведем некоторые из критериев.

1. Смирнов Е.С. “Таксономическая проблема заключается в “индикации”: от бесконечно большого числа признаков нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки”;
2. Рутковский Л. “Чем в большем числе существенных признаков сходны сравниваемые предметы, тем вероятнее их одинаковость и в других отношениях”.
3. Кожара В.Л. [7, с.49] “устойчивость таксономических структур (классификации – Е.Е.) зависит от числа оснований N (признаков – Е.Е.), чем больше, тем устойчивее. С увеличением N устойчивость ... повышается ... так, что при ... достаточно большом N , далее почти не повышается.”
4. Витяев Е.Е. [2] «Разбиение на классы должно производиться так, чтобы объекты одного класса подчинялись

одним и тем же закономерностям, объекты разных классов подчинялись разным группам закономерностей. Объекты одного класса, кроме того, должны обладать некоторой целостностью. Целостность определим как взаимную согласованность закономерностей каждой группы по предсказанию различных свойств объектов».

Принцип 3 является «принципом таксономической насыщенности», обнаруженный Кожара В.Л. [7]. В «естественной» классификации всегда существует такое N , что классификацию достаточно провести по любым N признакам и результат классификации от этого практически не изменится.

Представим эти критерии более формально. Объекты являются целостными образованиями, соединяющими в себе понятия, величины и законы рассматриваемой предметной области (ПО). Если законы рассматривать как аксиомы ПО, то каждый объект должен удовлетворять этим аксиомам с некоторой достоверностью (см. формальные определения в [3-5]). Поэтому объекты являются, в определённом смысле, моделями законов ПО.

Определим закономерную модель $M_{\mathfrak{C}} = \langle \Omega_{\mathfrak{C}}, Z_{\mathfrak{C}} \rangle$ класса \mathfrak{C} , где $\Omega_{\mathfrak{C}}$ – множество значений признаков, характеристик, величин $x_1, x_2, \dots, x_k, \dots$, принимающих на каждом объекте \mathbf{a} класса определенные значения (истинности, номинальные, числовые ...); $Z_{\mathfrak{C}}$ – закономерности вида $(x_1^i = y_{j_1}^i) \& (x_2^i = y_{j_2}^i) \& \dots \& (x_k^i = y_{j_k}^i) \Rightarrow (x_0^i = y_{j_0}^i)$ применимые к каждому объекту \mathbf{a} класса \mathfrak{C} и предсказывающие на нём по значениям $y_{j_1}^i, y_{j_2}^i, \dots, y_{j_k}^i$ величин $x_1^i, x_2^i, \dots, x_k^i$ значение $y_{j_0}^i$ величины x_0^i .

По критерию Е.С. Смирнова проблема индикации состоит в нахождении таких

«порождающих» признаков (см. точное определение в [2-3,10]), по значениям которых предсказываются все остальные значения признаков объектов класса (предполагается, что признаки не случайны и принадлежат онтологии ПО – системе понятий, фиксирующей предмет исследования).

Закон «естественной» классификации заключается в том, что в «естественных» классах значения «порождающих» признаков предсказывают потенциально бесконечное множество значений других признаков, что говорит об их глубинной взаимосвязи.

Из критерия Рутковского следует, что чем больше значений признаков фиксировать, тем вероятнее можно предсказать значения других признаков.

Различных «порождающих» наборов признаков огромное множество. Из принципа таксономической насыщенности Кожара В.Л. следует, что начиная с некоторого N любой набор признаков, содержащий более N признаков, будет порождающим. Таким образом, множество закономерностей $Z_{\mathcal{C}}$ должно быть огромно, что бы позволить предсказывать из значений порождающих признаков значения всех остальных признаков. Такое множество закономерностей осуществляет взаимопредсказание признаков друг друга. Это даёт определённый *системный закон взаимосвязи значений признаков*.

Объекты каждого класса характеризуются некоторой целостностью, которая обеспечивается особым способом организации взаимодействующих частей: "Между частями органичного целого ... существует не простая функциональная зависимость, а значительно более сложная система разнокачественных связей - структурных, генетических, связей субординации, управления и т.п., ... Взаимосвязь частей такова, что она выступает не в виде линейного причинного ряда, а в виде своеобразного *замкнутого круга* (выд. Е.Е.), внутри которого каждый элемент связи является условием другого и обусловлен им" ([9], статья «часть и целое»). Это определение целого согласуется с критерием 4 «естественности» классификации.

Целостность определяется, таким образом, системным законом взаимосвязи значений признаков, в котором значения при-

знаков взаимопредсказывают друг друга по «замкнутому кругу», что обеспечивает надёжность, непротиворечивость и согласованность взаимных предсказаний, образующих своеобразный *резонанс закономерных связей*. Резонанс состоит в том, что именно те значения признаков, которые характеризуют класс, взаимопредсказывают друг друга. Подобный резонанс, но в терминах нейронных сетей был обнаружен в рамках направления *adaptive resonance theory* [10].

В методах интеллектуального анализа данных (Knowledge Discovery in Data Bases and Data mining – KDD&DM) и машинного обучения (Machine Learning – ML) наличие слишком большого числа признаков в данных делает практически невозможным надёжное решение задач распознавания и предсказания. Это называется «проклятием размерности» – чем больше признаков, тем менее надёжным получается распознавание и предсказание. Закономерности, обнаруживаемые методами KDD&DM/ML (классификационные правила, решающие правила, решающие деревья, аппроксимации и т.д.), не могут дать адекватную оценку качества распознавания или предсказания, потому что их слишком много и возникает проблема переобучения, которую эти методы учесть не могут. «Проклятие размерности» возникает потому, что методы KDD&DM/ML не используют закон «естественной» классификации.

В «естественной» классификации «проклятие размерности» не возникает. Наоборот, чем больше признаков – тем она надёжнее. Распознавание и предсказание в «естественной» классификации осуществляется только после проведения классификации, использующей свой особый критерий резонанса закономерностей (см. [2-3,10]), и выделения групп закономерностей, которым удовлетворяют классы, в соответствии с критерием 4. Далее распознавание и предсказание осуществляется *только по тем закономерностям*, которые входят в описание класса и отражают закономерности объектов именно этого класса. Решающим в «естественной» классификации является её особый критерий резонанса закономерностей, определяющий согласованность закономерностей по взаимопредсказанию значений признаков.

Таким образом, решающим является критерий, оценивающий качество полученного резонанса и взаимопредсказаний значений признаков класса и, соответственно, надежность, непротиворечивость и согласованность предсказаний.

Методы и полученные результаты

Опираясь на критику П.К. Анохиным гипотезы суммации возбуждений на уровне нейрона [1], мы ввели другое определение формальной модели нейрона [4-5]. Это определение функции нейрона соответствует работам по замыканию условных связей на уровне нейрона. Оно также соответствует тому факту, что скорость ответа нейрона на условный сигнал, тем выше, чем больше условная вероятность ответа нейрона. Эта формальная модель нейрона была нами использована для моделирования функциональных систем П.К. Анохина [4]. На основании этой модели была разработана логическая модель адаптивной системы управления виртуальным аниматором, решающая задачу фуражирования [5]. Анимат показал хорошие результаты по скорости обучения, которые оказались лучше, чем у аниматоров, основанных на нейронных сетях.

В этих работах формальная модель клеточного ансамбля не вводится. Объем тезисов не позволяет её формально определить, но она в точности соответствует методу и алгоритму «естественной» классификации определённым в [2-3,11-13] и реализованным в [11-13].

Результаты применения разработанных систем «естественной» классификации DNANatClass [11] и GeneNatClass [12] в биоинформатике, показали её большую устойчивость к ошибкам второго рода (неверному распознаванию контрольных объектов). При решении задачи распознавания сайтов связывания транскрипционных факторов в работе [11] была достигнута ошибка второго рода менее 10^{-5} .

Работа клеточного ансамбля, основанного на «естественной» классификации действительности, лучше всего проявляется в восприятии.

Процесс восприятия направляется и организуется «образом мира» [8]. «Образ

мира» непрерывно во времени предвосхищает (апперцепцией) новые стимулы, возникающие после совершения перцептивных действий, и сравнивает их с реально полученными стимулами. Восприятие является процессом активного движения от «образа мира» к внешнему миру – непрерывным во времени и множественным по признакам процессом проверки адекватности «образа мира» на реальных стимулах. Только, если все многочисленные предсказания будут совпадать с реальными стимулами, только тогда есть восприятие и, в этом случае, «образ мира» работает адекватно, в противном случае он разрушается [8].

Такое функционирование «образа мира» можно объяснить, если предположить что «образ мира» включает «естественную» классификацию действительности. «Образ мира» в каждом акте предвосхищения стимулов, использует порождающий набор стимулов своей внутренней «естественной» классификации. Этот порождающий набор включает стимулы текущей ситуации и предполагаемого действия и динамически (когда нейроны, активирующие действие, уже сработали, а само действие ещё не совершилось) предсказывает стимулы той части образа, которая появится после совершения действия.

Резонанс закономерных связей «естественной» классификации «образа мира» работает динамически – стимулы текущей ситуации плюс стимулы от совершаемого действия предвосхищают ожидаемые стимулы от другой части образа, которая будет воспринята после совершения перцептивного действия. Резонанс – совпадение предвосхищаемых стимулов и реально поступивших – возникнет в результате восприятия поступившей стимуляции и её совпадения с предсказанной. После этого взгляд переместится в другое место образа и снова осуществится резонанс в результате восприятия совпадения предвосхищенных и реальных стимулов. Непрерывные саккадические движения глаз поэтому нужны для осуществления резонанса и восприятия.

Выводы

Функция клеточного ансамбля очень важна для понимания работы мозга. На

клеточном ансамбле лежит основная роль в настройке на реальность. Он осуществляет непрерывный и точный по предсказанию контакт с реальностью, постоянно проверяющийся на адекватность.

Поэтому возникает задача всесторонней экспериментальной проверки гипотезы о том, что: функция клеточного ансамбля состоит в «естественной» классификация действительности резонансом закономерных связей.

Список литературы

1. Анохин П.К. Системный анализ интегративной деятельности нейрона // П.К. Анохин. Очерки по физиологии функциональных систем. М.: Медицина, 1975. с. 444.
2. Витяев, Е.Е. Классификация как выделение групп объектов, удовлетворяющих разным множествам согласованных закономерностей. // Анализ разнотипных данных (Выч. сист. 99), Новосибирск, 1983, с. 44-50.
3. Витяев Е.Е., Морозова Н.С., Сутягин А.С., Лапардин К.А. Естественная классификация и систематика как законы природы // Анализ структурных закономерностей (Вычислительные системы вып. 174), Новосибирск, 2005, с.80-92
4. Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова. Нейроинформатика, 2008, том 3, № 1, стр. 25-78
5. Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления. Нейроинформатика, 2008, том 3, № 1, стр. 79-107
6. Забродин, В.Ю. О критериях естественной классификации. – НТИ, сер.2, 1981, №8.
7. Кожара В.Л. Анализ информативно насыщенных таксономических структур как способ выявления географических закономерностей \\\ Дисс. к.геогр.н., М., 1989.
8. Смирнов С.Д. Психология образа: проблема активности психического отражения. МГУ, М., 1985, с.232.
9. ФЭС. Философский энциклопедический словарь. М. "Советская энциклопедия" 1989. 816с.
10. Carpenter, G.A. & Grossberg, S. (2003), Adaptive Resonance Theory, In Michael A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, Second Edition (pp. 87-90). Cambridge, MA: MIT Press
11. Vityaev E.E., Lapardin K.A., Khomicheva I.V., Proskura A.,L. Transcription factor binding site recognition by regularity matrices based on the natural classification method. Intelligent Data Analysis. Special issue: "New Methods in Bioinformatics Presented at the fifth International Conference on Bioinformatics of Genom Regulation and Structure" eds. Evgenii Vityaev and Nikolai Kolchanov. v.12(5), IOS Press, 2008 pp. 495-512
http://www.math.nsc.ru/AP/ScientificDiscovery/PDF/regularity_matrixes.pdf
12. Vityaev E.E., Kostin V.V., Podkolodny N.A., Kolchanov N.A. Natural classification of nucleotide sequences. // Proc. of the Third International Conference On Bioinformatics of Genome Regulation and Structure (BGRS'2002, Novosibirsk, Russia, July 14-20, 2002), v3, ICG, Novosibirsk, 2002, pp. 197-199
13. Scientific Discovery web site
<http://www.math.nsc.ru/AP/ScientificDiscovery>