



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Применение методов формальных понятий к построению оптимальной архитектуры нейронной сети

Аспирант 2-го года обучения департамента анализа данных и искусственного интеллекта ФКН НИУ  
ВШЭ (Москва)

**Махажанов Нуртас**

[makhazhanovn@gmail.com](mailto:makhazhanovn@gmail.com)

Научный руководитель:

д.ф.-м.н., профессор, зав. департамента анализа данных и искусственного интеллекта

**Кузнецов Сергей Олегович**

03.03.2016

# Методы анализа формальных понятий

Формальный контекст – это тройка  $K = (G; M; I)$ , где  $G$  – множество объектов,  $M$  – множество признаков, а бинарное отношение  $I \subseteq G \times M$  определяет, какой объект каким признаком обладает. Предикат  $gIm$  означает, что объект  $g$  имеет признак  $t$ . Для подмножеств множеств объектов и признаков  $A \subseteq G$  и  $B \subseteq M$  операторы Галуа определяются следующим образом:

$$A' = \{t \in M \mid gIm \text{ для всех } g \in A\},$$

$$B' = \{g \in G \mid gIm \text{ для всех } t \in B\},$$

Оператор  $''$  (применение оператора  $'$  дважды) называется *оператором замыкания*. Множество объектов  $A \subseteq G$ , таких что  $A'' = A$ , называется *замкнутым*.

Пара  $(A, B)$ , такая что  $A \subseteq G, B \subseteq M, A' = B$  и  $B' = A$ , называется *формальным понятием* контекста  $K$ . Множества  $A$  и  $B$  замкнуты и называются *объемом* и *содержанием* формального понятия  $(A, B)$  соответственно. Для множества объектов  $A$  множество их общих признаков  $A'$  определяет сходство объектов множества  $A$ , а замкнутое множество  $A''$  есть множество схожих объектов (с общими признаками из  $A'$ ).

# Пример контекста

$G \setminus M$	w	y	g	b	f	$\bar{f}$	s	$\bar{s}$	r	$\bar{r}$	фрукт
яблоко		×				×	×		×		+
грейпфрут		×				×		×	×		+
киви			×			×		×		×	+
слива				×		×	×			×	+
кубик			×		×		×			×	-
яйцо	×				×		×			×	-
теннисный мяч	×					×		×	×		-

Сокращения:

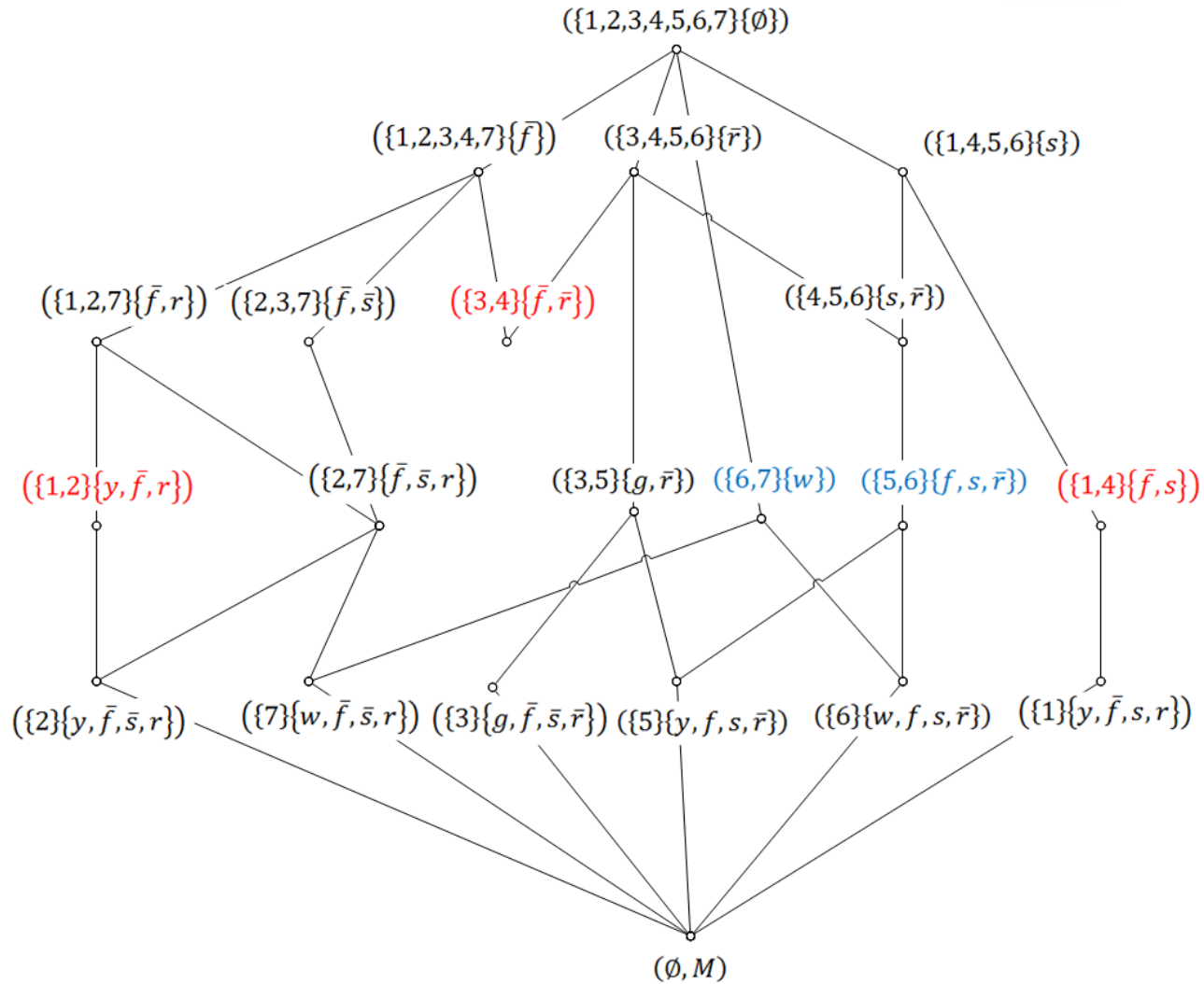
“g” – зеленый, “y” - желтый, “w” – белый, “b” - синий, “f” - твердый, “s” - гладкий, “r” - круглый

Общее число формальных понятий контекста  $K = (G; M; I)$  довольно существенно: в худшем случае  $2^{\min\{|G|, |M|\}}$ . Пусть  $K_+, K_-$  - положительный и отрицательный контексты относительно некоего целевого признака  $\omega$ , а

$$K_{\pm} := (G_+ \cup G_-, M \cup \{\omega\}, I_+ \cup I_- \cup G_+ \times \{\omega\})$$



# Решетка понятий



# Порождение понятий

В теории формальных понятий используется бинарное определение признаков. Поэтому для начала нам необходимо будет провести операцию шкалирования данных.

Шкалирование – это процесс приведения первичной базы к такому виду, когда все признаки представлены в двоичном виде.

Для всех числовых данных мы применим сравнение со средним значением характеристики по всем пациентам. То есть, если средняя величина характеристики «Диаметр опухоли» равна  $\bar{D}$ , то все пациенты с характеристикой выше  $\bar{D}$  будут считаться как обладающие признаком, иначе как не обладающие.

Главная проблема при применении АФП экспоненциальный рост времени работы алгоритмов. Для обработки контекста с множеством бинарных отношений / потребуется выполнение  $2^I$  операций. Поэтому мы будем порождать лишь общие понятия (для данных в 500 объектов и 30 атрибутов, будем порождать понятия, состоящие максимум из четырех признаков).



# Порождение понятий

Для порождения понятий будем использовать алгоритм *AddExtent*. Сначала набор понятий  $H$  составляется из всех признаков  $M$ . Далее итеративно добавляем к понятию  $h_i$  признаки из  $M$ .

*Псевдокод порождения новых понятий:*

```
lenPrev = length(M)
for i = 0; i < latticeDepth do
  lenCurr = length(H)
  for j = lenPrev+1; j < lenCurr do
    for k = 0; k < length(m) do
       $h^{new} \leftarrow (h_j \cup m_k)$ 
      if  $((h^{new})' \neq \emptyset)$  and  $(h^{new} \in H)$  then
        H append  $h^{new}$ 
      end if
    end for
  end for
  lenPrev = lenCurr
end for
```

# Отбор наилучших гипотез

После того как мы сформировали все возможные комбинации понятий, нам необходимо отобрать набор положительных и отрицательных гипотез для дальнейшего внедрения их в нейронную сеть. Общий балл гипотезы  $h_i$  будет основан как на достоверности, так и коэффициенте поддержки:

$$Score(h_i) = \alpha Confidence(h_i) + (1 - \alpha) Support(h_i)$$

Мера достоверности предсказаний – это соотношение количества правильно предсказанных объектов к общему количеству объектов, где встречалась данная гипотеза:

$$Confidence(h_i) = \frac{h'_i \cap T}{h'_i}$$

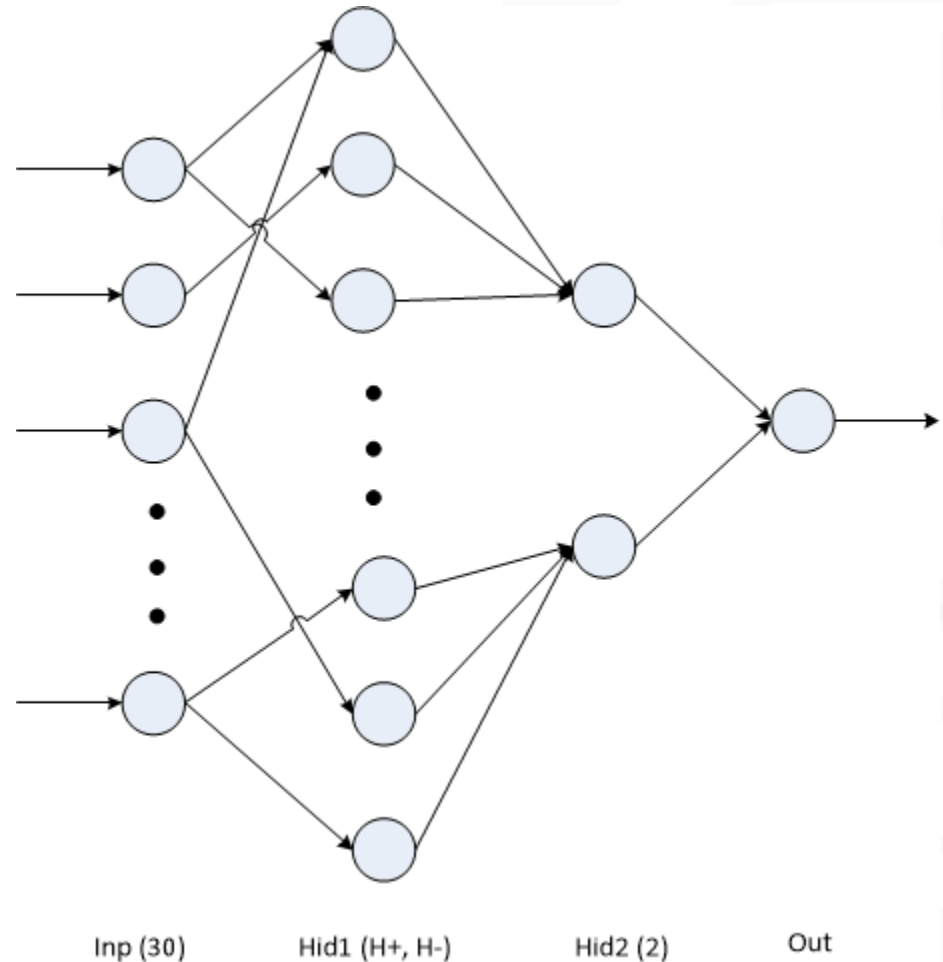
где  $h'_i = \{g \in G | gIm \text{ для всех } m \in h_i\}$  - это мощность множества, покрываемого гипотезой  $h_i$ .

Поддержка гипотезы:

$$Support(h_i) = \frac{h'_i}{G}$$

Сеть будет состоять из трех слоев:

- Входной слой  $Inp$  будет состоять из всех 30 признаков  $t \in M$ .
- Первый скрытый слой  $Hid_1$ , состоящий из таких комбинаций признаков, что каждый нейрон образует одну из отобранных гипотез.
- Второй скрытый слой  $Hid_2$  будет состоять из двух нейронов, каждый из которых будет прогнозировать, к какому классу принадлежит объект.





# Архитектура нейронной сети

- Все выходы нейронов слоя  $Hid_1$ , ассоциированных с положительными гипотезами  $hid^+$  с нейроном скрытого слоя  $Hid_2$ , ассоциированных с положительным классом. Соответственно, все нейроны, относящиеся к отрицательным гипотезам  $hid^-$  соединяются с нейроном, ассоциированным с отрицательным классом.
- Меры Precision и Recall наиболее полно характеризуют модель. Но часто бывает так, что увеличение одной из этих метрик приводит к уменьшению другой. Поэтому бывает удобно для общей характеристики модели использовать одну величину, учитывающую как Precision, так и Recall. Для этой цели мы используем метрику  $F_1$ :

$$F_1 = \frac{Precision * Recall}{Precision + Recall}$$

Таблица 1. Число генерируемых гипотез и время выполнения в зависимости от числа уровней верхней полурешетки.

N	Количество генерируемых гипотез		Оценочное время выполнения	Количество принятых гипотез	
	positive	negative		positive	negative
1	31	31	2 сек	5	6
2	490	490	20 сек	18	14
3	4617	4608	2 мин	19	15
4	28186	29615	20 мин	21	30
5	103361	131247	6 часов	19	19

Таблица 2. Характеристики качества модели в зависимости от числа уровней верхней полурешетки.

N	Параметры качества предсказания FCA без NN				Параметры качества предсказания FCA + NN			
	Accuracy	Precision	Recall	$F_1$	Accuracy	Precision	Recall	$F_1$
1	92.8	76.5	92.9	83.9	92.8	100.0	77.3	62.5
2	82.6	58.8	66.7	62.5	95.7	88.2	93.8	90.9
3	82.6	58.8	66.7	62.5	89.9	94.1	72.7	67.2
4	82.6	58.8	66.7	62.5	89.9	94.1	72.7	67.2
5	82.6	58.8	66.7	62.5	95.7	82.4	100	90.4

1. Shavlik, W.J., Towell, G.G.: Kbnns: Knowledge based artificial neural networks. *Artificial Intelligence* (70), 119–165 (1994).
2. Tsopez, N., Mephu Nguifo, E., Tindo, G.: CLANN: Concept-Lattices-based Artificial Neural Networks. *Proceedings of fifth Intl. Conf. on Concept Lattices and Applications (CLA 2007)*, Montpellier, France, October 24-26, 2007, pp. 157–168.
3. Endres, Dominik, and Földiák, Peter: Interpreting the neural code with Formal Concept Analysis. *Advances in Neural Information Processing Systems 21*, pp 425–432, Eds: Koller, D., Schuurmans, D., Bengio, Y., and Bottou, L., MIT Press, Cambridge, MA, 2009.
4. Engelbert Mephu Nguifo, Norbert Tsopez, Gilbert Tindo. Multiclass Concept Lattice-Based Artificial Neural Network. *Constructive Neural Networks // Studies in Computational Intelligence*. Vol. 258, 2009, pp 103-121.
5. Sebastian Rudolph, Using FCA for Encoding Closure Operators into Neural Networks//Conceptual Structures: Knowledge Architectures for Smart Applications. 15th International Conference on Conceptual Structures, ICCS 2007. Sheffield, UK, July 22-27, 2007, pp 321-332.
6. Cherukuri Aswani Kumar, M. S. Ishwarya, Chu Kiong Loo. Modeling Associative Memories Using Formal Concept Analysis. *Computational Intelligence in Information Systems. Advances in Intelligent Systems and Computing*. Vol. 331, 2015, pp 109-118.

7. Шашкин Л.О. Применение методов эволюционного моделирования для оптимизации множества ДСМ-гипотез. // Искусственный интеллект и принятие решений. –Март 2010, С. 33-39.
8. Eugene M. Norris: Maximal Rectangular Relations. University of South Carolina, Columbia, SC 29208/USA.
9. Виноградов Д.В. ВКФ-метод порождения гипотез: программная реализация // Труды 14-ой Национальной Конференции по Искусственному Интеллекту (КИИ-14). – 2014. – Том 2. – С. 252-258.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Спасибо!

**Махажанов Нуртас**

Аспирант 2-го года обучения департамента анализа данных и искусственного интеллекта ФКН НИУ ВШЭ

[makhazhanovn@gmail.com](mailto:makhazhanovn@gmail.com)