



Министерство науки и высшего образования Российской
Федерации Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Московский государственный технический университет имени
Н.Э. Баумана (национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ Информатика и системы управления

КАФЕДРА Системы обработки информации и управления

Рубежный контроль №2

По курсу

«Методы машинного обучения в АСОИУ»

Выполнил:

ИУ5-22М Бабин А.С.

20.04.2024

Проверил:

Гапанюк Ю.Е.

2024 г.

Представление решётки понятий АФП

Анализ формальных понятий АФП (от англ. Formal Concept Analysis – FCA) это теория анализа данных, которая идентифицирует концептуальные структуры в наборах данных и может представлять их в виде решётки понятий.

Концептуальная решетка понятий состоит из набора формальных понятий из формального контекста и отношений между ними. Формальные объекты принято отмечать на графе чуть ниже от узла, а формальные атрибуты - чуть выше узлов, к которым они относятся.

Самый верхний и самый нижний узлы в решетке понятий являются особыми. Верхний узел содержит все формальные объекты (объём формальных понятий). Нижняя концепция содержит все формальные атрибуты (содержание формальных понятий).

Чтобы найти объём формальной концепции, нужно проследить все пути, которые ведут вниз от узла и выделить все попавшиеся на этом пути формальные объекты.

Чтобы восстановить содержание формального понятия, нужно проследить все пути, которые ведут снизу вверх от узла и выделить все встреченные формальные атрибуты.

Решётка понятий основана по понятии иерархии. Есть более общие формальные понятия и более частные, то есть вложенные формальные понятия. Отношение “подпонятие – понятие” является транзитивным, что означает, что подпонятие является частным по отношению к любому понятию, которого можно достичь, двигаясь от него вверх.

Если формальное понятие имеет атрибут, то этот атрибут наследуется всеми его подпонятиями. Это соответствует понятию “наследование”, используемому в библиотеках классов объектно-ориентированного моделирования.

По этой причине АФП подходит для моделирования и анализа объектно-ориентированных библиотек классов.

Примеры формального контекста и соответствующей решётки понятий представлены на рисунках 1-2.

	cartoon	real	tortoise	dog	cat	mammal
Garfield	X				X	X
Snoopy	X			X		X
Socks		X		X	X	X
Greyfriar's Bobby		X		X		X
Harriet		X	X			

Рисунок 1 – формальный контекст

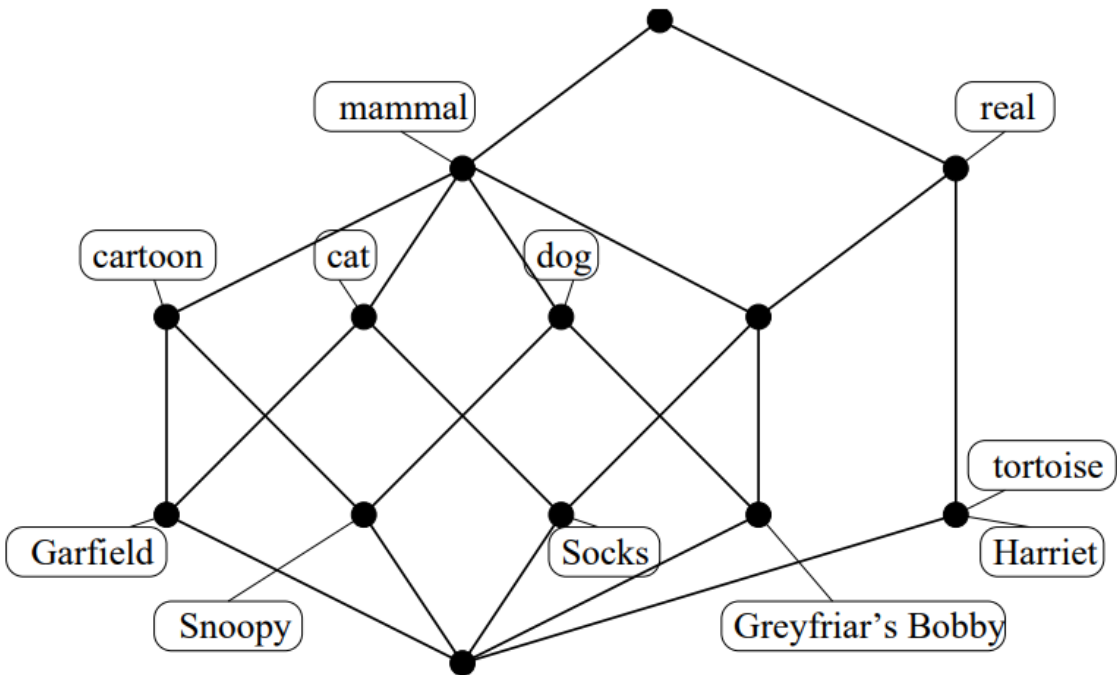


Рисунок 2 – решётка понятий

Довольно часто размер решетки велик, а ее структура сложна. Было предложено много методов: разделять набор элементов и получать вложенные линейные диаграммы или раскладывать решетку, используя соответствия или допуски, чтобы получить представления факторов или атласа.

Другие предложения сосредоточены на верхней части решетки (решетки айсберга) или при увеличении / уменьшении масштаба с использованием таксономии is-a по атрибутам (градуированные/обобщенные атрибуты).

Методы визуализации являются ценными инструментами для интерпретации извлеченных знаний и становятся все более важными, особенно в контексте больших и сложных данных.

Многозначный контекст

Часто данные напрямую не кодируются в двоичной форме. Если атрибут может принимать не только значение присутствует/отсутствует, то требуется дополнительная работа с такими данными.

Многозначный контекст – это контекст $K = (G, M, W, I)$, такой, что:

G - это набор объектов;

M - набор имен многозначных атрибутов;

W - набор значений атрибутов;

$I \subseteq G \times M \times W$.

Многозначный классификационный контекст может быть преобразован в бинарные формальные контексты с помощью алгоритма масштабирования формальных понятий [9]. Шкала для атрибута m из (G, M, W, I) представляет собой бинарный контекст $S_m := (G_m, M_m, I_m)$ такой, что $m(G) \subseteq G_m$.

На практике набор объектов остается неизменным; имя каждого атрибута m заменяется масштабируемыми атрибутами $s_m \in M_m$. Выбор подходящего набора масштабов зависит от интерпретации и обычно выполняется с помощью эксперта в предметной области. Номинальная, порядковая, интервальная и дихотомическая шкалы являются примерами элементарных шкал.

Альтернативным способом обработки таких атрибутов без масштабирования является использование структур шаблонов. Структуры шаблонов состоят из объектов с описаниями, которые допускают работу с ними в виде полурешетки, и естественным образом возникают из упорядоченных данных.

Импликации

Важной областью в математическом аппарате метода анализа формальных понятий, которую необходимо упомянуть, является использование импликаций. Импликация показывает зависимости между формальными атрибутами.

Импликацией формального контекста $K = (G, M, I)$ в анализе формальных понятий называется признаковая зависимость вида $A \rightarrow B$, где $A, B \subseteq M$, при условии, что все объекты, обладающие A , также обладают всеми признаками из B , т.е. $A' \subseteq B'$.

Импликация в АФП является частным случаем такой признаковой зависимости как ассоциативные правила (которые также называются частными импликациями) с достоверностью (confidence), равной единице. Частичная импликация в контексте K отличается тем, что не располагает полной поддержкой в обучающей выборке объектов. Вводимый в анализ формальных понятий показатель достоверности частичных импликаций позволяет, субъективно выбрав порог достоверности, расширить множество релевантных эмпирических закономерностей, но, как и прежде, не сопровождается использованием дополнительных данных о рассматриваемой предметной области [8].

А связь импликаций и функциональных зависимостей позволяет использовать так называемый базис импликаций Дюкена-Гига для компактного представления функциональных зависимостей в виде их ограниченного множества, из которого все оставшиеся функциональные зависимости данного многозначного контекста выводимы по правилам Армстронга.

Таким образом, импликации могут быть использованы для поэтапного компьютерного построения концептуальных знаний, называемой исследованием атрибутов.