

Общий подход к задачам многопризнаковой классификации на основе парадигмы вербального анализа решений

Е. М. Фуремс

Федеральный исследовательский центр «Информатика и управление» РАН, Москва, Россия

Аннотация. Предложен общий подход к решению задач порядковой, номинально-порядковой и номинальной классификации на основе парадигмы вербального анализа решений, реализованный в методе STEPCLASS. Новый подход включает такие аспекты, как базовая структуризация задачи с определением ее типа, гибкая стратегия диалога с экспертом для выявления полного и согласованного набора классифицирующих правил с возможностью расширения структуры задачи, декомпозиция задач большой размерности с последующей агрегацией правил на множестве всех признаков.

Ключевые слова: вербальный анализ решений, многопризнаковая классификация, структуризация проблемной области, полнота и согласованность правил классификации, задачи классификации с иерархической структурой и/или большой размерности.

DOI 10.14357/20718594190406

Введение

В теории принятия решений задача классификации традиционно определяется как отнесение объектов, описываемых оценками по многим критериям, к предварительно заданным классам на основе информации, полученной от лица, принимающего решения (ЛПР).

Подавляющее большинство подходов, включая, помимо прочего, теорию полезности [1-3], методы относительного превосходства [4-6], теорию приближенных множеств [7-9] и ее расширение, основанное на доминировании и согласованности переменных [10], а также теорию размытых множеств [11, 12], предназначены для многокритериальной классификации с предварительно заданными классами, упорядоченными в соответствии с предпочтениями ЛПР (многокритериальная порядковая классификация или многокритериальная сортировка).

Как подчеркивается в [13], различие между задачами многокритериальной классификации и традиционными задачами классификации, решаемыми статистическими методами и методами распознавания образов, заключается в получении информации о предпочтениях ЛПР и в отношении оценок объектов по критериям, и в отношении заданных классов. Соответственно, такая классификация носит порядковый, а не номинальный характер. Тем не менее, в теории принятия решений при многих критериях также рассматривается задача многокритериальной номинальной классификации, где классы не могут быть упорядочены. Однако число подходов к ее решению относительно мало, например, метод многокритериальной фильтрации [14], методы PROAFTN [15], TRINOMFC [16], PROCFTN [17], процедуры

✉ Фуремс Евгения Марковна E-mail: fem@mail.ru

k-ближайшего сходства [18], а также методы NEXClass [19, 20], SMAA-Classification [21] и Cat-SD [22].

Во всех вышеперечисленных методах под классом понимается набор объектов, которые имеют большее сходство между собой, чем с объектами других классов. Классы предварительно задаются либо наверняка принадлежащими к ним объектами, либо границами, то есть наименее и наиболее предпочтительными объектами каждого класса, в случае многокритериальной порядковой классификации, или наименее и наиболее более типичными для них объектами в случае многокритериальной номинальной классификации. Соответственно, такие методы в той или иной степени осуществляют классификацию объектов на основе их сходства с использованием контролируемого обучения, требующего информации от ЛПР. В большинстве вышеназванных методов информацию, получаемую от ЛПР, называют предпочтениями, в том числе и для задачи номинальной классификации (например, в [23] отмечается, что результирующая классификация должна соответствовать предпочтениям ЛПР), и только за редким исключением – знаниями.

Однако в некоторых приложениях, например, построение баз знаний для систем поддержки диагностических решений, которые могут быть сведены к задаче классификации, должны использоваться не предпочтения ЛПР, а только его знания [24–26]. Действительно, даже в задаче порядковой классификации классы могут быть упорядочены не в соответствии с предпочтениями ЛПР, а по степени выраженности некоторого общего для объектов свойства, например, по тяжести одного и того же заболевания или по степени уверенности в его наличии. Соответственно, в таком контексте разумно заменить термин «критерий» термином «признак» и называть задачу классификации, основанную на знаниях, задачей многопризнаковой классификации. Кроме того, в такой задаче правильнее использовать вместо термина «ЛПР» термин «эксперт», понимая под этим лицо, которое принимает решения на основе своих знаний. Заметим, что термин «эксперт» применим также в задачах, основанных на предпочтениях, поскольку предпочтения ЛПР в той или иной степени тоже основаны на его знаниях. Далее, для общности изложения, если это не противоречит контексту, термины

«эксперт», «признак» и «знания» используются вместо терминов «ЛПР», «критерий» и «предпочтения», соответственно. Такая замена терминов носит исключительно концептуальный характер и не препятствует применению подходов к задаче многокритериальной классификации для задачи многопризнаковой классификации.

Все перечисленные выше подходы мотивированы до некоторой степени ограниченной возможностью непосредственного участия эксперта в решении задачи, что объясняется, в частности, существенными затратами времени на выявление его знаний или его нежеланием делиться своими знаниями. Поэтому в таких подходах основное внимание уделяется не организации эффективного диалога с экспертом для выявления его знаний, а способам вывода классификационных правил на основе их некоторой выборки, полученной из тех или иных источников (включая эксперта). Математический аппарат, используемый для такого вывода, делает результаты решения задачи классификации труднообъяснимыми (если вообще объяснимыми) не только для их конечных пользователей, но и для самого эксперта, что, естественно, влияет на доверие к ним. Кроме того, как подчеркивается в [27], многие известные методы классификации плохо справляются (если вообще справляются) с решением задач большой размерности: чем больше число признаков и их значений, тем с меньшей уверенностью можно говорить о достижимости полноты классифицирующих правил и точности выводимых из них заключений.

Иной подход используется в методах, основанных на парадигме вербального анализа решений (ВАР) [28, 29], предназначенных для задачи многопризнаковой классификации, когда есть эксперт, способный и желающий поделиться своими знаниями и, таким образом, участвовать на всех этапах решения задачи. Методы принятия решений, основанные на ВАР, используют информацию, получаемую от эксперта только в качественной (вербальной) форме, и обрабатывают ее без каких-либо количественных преобразований, а только на основе логических выводов, понятных и самому эксперту, и пользователю. Эти методы включают различные средства для уменьшения когнитивной сложности для эксперта и затрат времени на выявление его знаний.

1. Методы решения задач классификации на основе парадигмы ВАР

На основе парадигмы ВАР разработан ряд методов решения разных типов задач классификации: методы CLASS [30, 31], CYCLE [32] и DIFCLASS [33] для задачи порядковой классификации; метод решения, впервые поставленной в [25] задачи номинально-порядковой классификации, впоследствии получивший название NORClass [26]; метод STEPCLASS [34], исходно предназначенный для номинальной классификации.

Все эти методы основаны на том соображении, что, хотя эксперт может перечислить свои правила, их набор, почти наверняка, будет неполным и применимым только к типичным для каждого класса объектам. Это объясняется, помимо прочего, тем, что в своей повседневной деятельности эксперт не формулирует правила, а применяет их, анализируя конкретные объекты. В связи с этим, процедура выявления правил организована посредством моделирования объектов, подлежащих классификации, в виде комбинаций значений признаков и предъявления их эксперту для анализа и отнесения к соответствующему классу. Таким образом, класс – это категория, принадлежность объектов к которой либо непосредственно определена экспертом, либо выведена путем логических заключений из классифицирующих правил, сформулированных им для других объектов.

Эти методы реализуют различные подходы к сокращению числа обращений к эксперту для непосредственной классификации объектов, и, при этом, позволяют строить полные (с точностью до знаний эксперта) и согласованные базы классифицирующих правил даже для задач большой размерности.

Однако во всех этих методах, кроме STEPCLASS, отсутствует формализованный подход к структуризации задачи классификации, которая, по своей природе является слабоструктурированной или даже неструктурированной [35]. Наименования классов, признаков и их значений предполагаются предварительно и окончательно заданными экспертом и не могут быть дополнены в процессе классификации. Соответственно, эксперт, определяя, что объект не принадлежит ни к одному из заданных им

классов, вынужден отнести его в «безымянный» класс, даже если знает для него конкретный класс, который следовало бы добавить к набору исходных классов. Если эксперту не хватает информации для классификации объекта, он не имеет возможности ввести новые признаки (и их значения). Кроме того, для выявления правил эксперту одновременно показываются значения всех признаков данного объекта. Однако нередко эксперту достаточно знать значения только части признаков, чтобы принять решение о принадлежности тому или иному классу при любых значениях остальных признаков. Такой подход не только увеличивает когнитивную нагрузку на эксперта при анализе объекта и его классификации, но и приводит к увеличению числа обращений к нему.

Как уже отмечалось выше, метод STEPCLASS исходно был разработан для номинальной многопризнаковой классификации, где и классы, и значения признаков не могут быть упорядочены. В [36, 37] показывается, как метод STEPCLASS может быть использован для многопризнаковой порядковой классификации, а в [38, 39] описывается его расширение для номинально-порядковой классификации (классы не упорядочены, но значения некоторых, по крайней мере, признаков могут быть упорядочены по типичности для каждого класса).

Далее описывается общий подход к решению задач классификации любого из вышеуказанных типов, реализованный в расширенной версии метода STEPCLASS, в которой сохраняются все достоинства вышеперечисленных методов, основанных на парадигме ВАР, и преодолеваются присущие им недостатки. Показывается, как такой общий подход позволяет уменьшить когнитивную нагрузку на эксперта, существенно сократить число объектов, предъявляемых ему для непосредственной классификации, всесторонне проверить согласованность его правил и в результате получить их полную и согласованную совокупность. Кроме того, метод STEPCLASS можно использовать для задач классификации с исходно иерархической структурой и задач классификации большой размерности, преобразуемых в задачи с иерархической структурой посредством их декомпозиции по признакам и/или по классам на подзадачи меньшей размерности [40].

2. Постановка и структуризация задачи многопризнаковой классификации

2.1. Определение общей структуры и типа задачи классификации

В методе STEPCLASS задача классификации ставится в виде двух взаимосвязанных подзадач: (1) структуризация, то есть предварительное определение наименований классов, признаков и их значений, а также определение типа классификации и (2) выявление классифицирующих правил в рамках такой структуры с возможностью ее расширения [26]. В общем случае структура задачи определяется следующим образом:

$$S = (C, Q, \{Q_m\}, \{V_m\}),$$

где $C = \{C_1, C_2, \dots, C_L\}$; $Q = \{Q_1, Q_2, \dots, Q_M\}$ и $Q_m = \{q_{m1}, q_{m2}, \dots, q_{mn_m}\}$ – наборы наименований классов, признаков и их значений, соответственно, подлежащих заданию экспертом; $V_m \equiv [v_{mj}^l]$ – $L \times n_m$ матрица допустимости значений признака Q_m для классов: $v_{mj}^l = 1$, если значение $q_{mj} \in Q_m$ допустимо для класса C_l , в противном случае, $v_{mj}^l = 0$, $m = \overline{1, M}$.

Декартово произведение шкал признаков $A = Q_1 \times Q_2 \times \dots \times Q_M$ определяет набор M -признаковых описаний всех гипотетически возможных объектов данной проблемной области; каждый объект $a_i \in A$ описывается кортежем $a_i = (a_{i1}, a_{i2}, \dots, a_{iM})$, где $a_{im} \in Q_m$, $m = \overline{1, M}$.

Примечание 1. Такие кортежи организованы в возрастающем лексикографическом порядке номеров значений признаков, причем и признаки, и их значения нумеруются в порядке, указанном экспертом, то есть $a_1 = (q_{11}, q_{21}, \dots, q_{M1})$, $a_2 = (q_{11}, q_{12}, \dots, q_{M2})$, и т.д.

Примечание 2. Наименования классов, признаков и их значений задаются на естественном языке.

Примечание 3. Помимо названных экспертом классов, вводятся два дополнительных класса, C^{Un} и C^{Inc} . Класс C^{Un} предназначен для объектов, которые эксперт не может отнести ни к одному классу из C . Класс C^{Inc} предназначен для объектов, чьи описания содержат несовместимые значения признаков (поскольку

набор A содержит описания всех гипотетически возможных объектов данной проблемной области, в том числе не существующих в реальности).

Задача ставится следующим образом. *Дано:* определенная проблемная область, каждый объект которой может принадлежать одному или нескольким классам. *Требуется:* разделить объекты из A на классы из C на основе знаний эксперта так, чтобы результирующая классификация была полна (с точностью до знаний эксперта) и согласована.

Примечание 4. В задаче порядковой классификации каждый объект из A может принадлежать только к одному классу.

Примечание 5. Полнота означает, что для любого объекта $a_i \in A$ $C_i^+ \cup C_i^- = C$, или $a_i \in C^{Un}$, или $a_i \in C^{Inc}$, где $C_i^+ \in C$ означает набор классов, к которым принадлежит объект a_i и $C_i^- \in C$ – набор классов, к которым a_i не принадлежит. Очевидно, что если $C_i^+ \neq \emptyset$, объект a_i не принадлежит ни к классу C^{Un} , ни к классу C^{Inc} .

Примечание 6. Согласованность означает, что для каждого объекта из A может быть задано любое число правил, но все они должны относить такой объект к одному и тому же набору классов.

2.2. Определение общей структуры и типа задачи классификации

В методе STEPCLASS предусмотрены два подхода к определению элементов структуры задачи классификации [16]: явная структуризация и структуризация на примерах, которые с помощью специальных наводящих вопросов позволяют определить перечисленные выше параметры структуры.

На этом этапе также определяется тип задачи: порядковая, номинально-порядковая или номинальная. Если эксперт может упорядочить классы по их предпочтительности или по степени выраженности некоторого общего свойства, решаемая задача является порядковой классификацией. На множестве классов строится соответствующее отношение линейного порядка $P^C = \{(C_i, C_j) \in C \times C \mid i < j; i, j = \overline{1, L}\}$, а на шкале каждого признака Q_m – соответству-

ющие отношения квазипорядка: $R_m = \{(q_{mi}, q_{mj}) \in Q_m \times Q_m\}$, $m = \overline{1, M}$, такие, что пара значений $(q_{mi}, q_{mj}) \in R_m$, если значение q_{mi} не менее предпочтительно или не менее типично для такого общего свойства, соответственно, чем значение q_{mj} .

Если классы не могут быть упорядочены (например, они соответствуют разным заболеваниям в задаче дифференциальной диагностики), но эксперт может упорядочить значения некоторых, по крайней мере, признаков по их типичности для каждого класса независимо от значений других признаков, решаемая задача является номинально-порядковой. На шкале каждого такого признака строятся бинарные отношения квазипорядка $R_m^l = \{(q_{mi}, q_{mj}) \in Q_m \times Q_m\}$, $m \in (1, \dots, M), l = \overline{1, L}$, так что пара $(q_{mi}, q_{mj}) \in R_m^l$, если значение q_{mi} не менее типично для класса C_l , чем значение q_{mj} .

Наконец, если эксперт не может упорядочить ни классы, ни шкалы признаков, решается задача номинальной классификации.

Поскольку, как отмечалось выше, структуризация задачи классификации не свойственна профессиональной деятельности эксперта, нельзя гарантировать, что он сформирует на этом этапе полные наборы наименований классов, признаков и их значений, и такая структура может считаться лишь предварительной. Метод STEPCLASS позволяет эксперту, в случае необходимости, расширить структуру задачи на этапе выявления правил.

2.3. Задачи классификации с иерархической структурой

Признаки, перечисленные экспертом в режиме непосредственной структуризации или выведенные из примеров, приведенных для каждого класса в режиме структуризации на примерах, могут быть «сложными». Это означает, что на практике значения таких признаков не проявляются у объектов, подлежащих классификации, непосредственно, а зависят от совокупности значений других признаков. В таком случае конечный пользователь будет располагать не значениями таких признаков, а некоторой базовой информацией, из которой такие значения должны

быть определены. Следовательно, в отношении каждого признака эксперту задается вопрос о том, как пользователь сможет определять значения признаков – непосредственно или на основе значений других признаков.

Если для какого-то признака такая информация нужна, эксперту предлагается применить одну из вышеупомянутых процедур структуризации, чтобы определить признаки и их возможные значения для классов, наименованиями которых являются значения такого признака. Тем самым, формируется подзадача для такого признака.

Например, в подробно описанной в [39] задаче оценки качества письменных переводов с иностранных языков одним из критериев эксперт назвал «Передача смысла» с тремя значениями: «смысл передан хорошо», «смысл передан удовлетворительно» и «смысл передан плохо», а на вопрос, как пользователь (редактор) сможет определить качество передачи смысла, эксперт ответил, что для этого необходимо оценить точность и полноту перевода. Для критерия «Точность перевода» были сформулированы две оценки: «Перевод отражает смысл оригинала», «Перевод не отражает смысл оригинала», а для критерия «Полнота» – три оценки: «Все элементы смысла переданы», «Небольшое число элементов смысла не передано» и «Много не переданных элементов смысла». В результате, для критерия «Передача смысла» была создана подзадача классификации с тремя классами (соответствующими его вышеуказанным значениям) и двумя критериями «Точность» и «Полнота».

Метод STEPCLASS предусматривает также возможность определения значений «сложных» признаков по заданным экспертом формулам, в которые должны подставляться «простые» признаки.

Например, при построении с использованием метода STEPCLASS базы знаний для системы поддержки принятия решений (СППР) «Оценка кредитоспособности предприятий», рассмотренной в [39], в качестве одного из признаков эксперт назвал «Коэффициент текущей ликвидности» ($K_{ТЛ}$) со следующими вербальными значениями: $K_{ТЛ} \geq 0.2$; $0.15 \leq K_{ТЛ} \leq 0.2$; $K_{ТЛ} < 0.15$ и ввел следующую формулу для расчета этого коэффициента:

$$K_{ТЛ} = \frac{A_1 + A_2 + A_3}{\Pi_1 + \Pi_2},$$

где A_1 – наиболее ликвидные активы (денежные средства предприятия и краткосрочные финансовые вложения); A_2 – быстро реализуемые активы (дебиторская задолженность и прочие активы); A_3 – медленно реализуемые активы (запасы (без статей «Прочие запасы и затраты» и «Расходы будущих периодов»), а также статьи из раздела I баланса «Долгосрочные финансовые вложения», уменьшенные на величину статьи «Вложения в уставные фонды других предприятий»); Π_1 – наиболее срочные пассивы: кредиторская задолженность, прочие пассивы, а также ссуды, не погашенные в срок; Π_2 – краткосрочные пассивы: краткосрочные кредиты и заемные средства.

Тем самым, для признака $K_{ТЛ}$ была сформирована подзадача с пятью вышеуказанными признаками и приведенной выше формулой в качестве решающего правила. После того как пользователь СППР введет конкретные значения соответствующих признаков этой подзадачи, значение $K_{ТЛ}$ будет вычислено по приведенной выше формуле и сопоставлено с вышеуказанными интервалами его значений.

Такие подзадачи строятся для каждого «сложного» признака, и в результате получается иерархическая структура, состоящая из основной задачи классификации и подзадач для ее признаков. Заметим, что подзадачи также могут включать «сложные» признаки, и в таком случае для них создаются свои подзадачи, и так далее.

2.4. Задачи классификации большой размерности

Эффективность существующих подходов к извлечению знаний страдает от так называемого «проклятия размерности» [41]. Как объясняется в [27], во многом это обусловлено наличием в структуре задачи избыточных и нерелевантных признаков.

Метод STEPCLASS свободен от избыточности признаков и отсутствия у них релевантности, поскольку признаки определяются самим экспертом на этапе структуризации (и, в случае необходимости, добавляются на этапе классификации), и, даже если эксперт задаст на этапе структуризации «лишние» признаки, он никогда не будет их использовать на этапе классификации.

Кроме того, благодаря особой организации диалога с экспертом для выявления его правил, метод STEPCLASS позволяет решать задачи большей размерности, чем вышеупомянутые методы CLASS, CYCLE, DIFCLASS и NORClass. Как уже отмечалось, во всех этих методах объект, подлежащий классификации, предъявляется эксперту в форме его описания значениями всех признаков одновременно, в результате чего ему приходится анализировать объект сразу с учетом всех признаков (в том числе и избыточных), а с другой стороны, требует предъявления эксперту большего числа объектов для непосредственной классификации. При применении метода STEPCLASS для каждого объекта также известны значения всех признаков, но, как показано ниже, эксперт сам запрашивает только те признаки, значения которых он хочет узнать. Тем не менее, если эксперт на этапе структуризации определяет большое число признаков и большое число значений каждого признака, число гипотетически возможных объектов, подлежащих классификации, может достигать нескольких миллионов или даже миллиардов (число таких объектов равно мощности декартова произведения шкал признаков). Таким образом, практически неосуществимо построить полную базу правил для такой задачи за приемлемое время, несмотря на то, что для построения полной базы правил непосредственно классифицируется лишь часть объектов. Однако на практике эксперт не работает с полным набором признаков одновременно. Значения некоторых из них он анализирует либо независимо от других, либо с учетом предварительного решения, принятого на основе анализа комбинаций значений других признаков. Это соображение реализуется в методе STEPCLASS посредством декомпозиции задачи большой размерности и сведения, тем самым, ее структуры в иерархическую структуру подзадач приемлемой размерности.

В STEPCLASS используется два способа декомпозиции задачи классификации: по признакам и классам. Декомпозиция по признакам выполняется следующим образом. Эксперт анализирует набор признаков, определенных им на этапе структуризации, чтобы разделить их на подгруппы, характеризующие разные аспекты, которые ему необходимо знать, чтобы классифицировать объекты. Другими словами, такие подгруппы формируются так, чтобы экс-

перт мог сформировать правила для классификации всех гипотетически возможных комбинаций значений признаков каждой подгруппы. В отличие от задач классификации, исходно имеющих иерархическую структуру в случае задач большой размерности, иерархическая структура подзадач строится не сверху вниз (от основной задачи к подзадачам), а снизу вверх (от подзадач к основной задаче). Наименования признаков и их значения, определенные экспертом на этапе структуризации, используются в подзадачах нижнего уровня. Наименованиями признаков верхнего уровня становятся те аспекты, для которых формировались соответствующие подзадачи, а значениями каждого такого признака — наименование классов соответствующей подзадачи.

Например, в задаче дифференциальной диагностики бронхиальной астмы у детей [42], врач-эксперт исходно сформировал на этапе структуризации набор из 34 признаков с 5-10 значениями у каждого. Этот набор признаков эксперт объединил в группы, соответствующие разделам унифицированной истории болезни пульмонологического больного: жалобы, осмотр (клинические данные), анамнез, данные клинического анализа крови, данные рентгенологического исследования и данные функции внешнего дыхания. Это позволило построить иерархическую структуру задачи, которая решалась экспертом, начиная с подзадач нижнего уровня. Сначала были выявлены классифицирующие правила для подзадачи «Жалобы» с семью признаками, в которой к исходно названным 13 классам добавлялись слова «Предварительный диагноз по жалобам». Поскольку метод STEPCLASS позволяет добавлять наименования классов на этапе классификации, а также в связи с тем, что признаков только этой подзадачи эксперту было явно недостаточно для дифференциальной диагностики большинства комбинаций значений признаков, эксперт мог относить такие комбинации сразу к нескольким заболеваниям (классам) с использованием союза «или» (каждая такая совокупность классов считалась новым классом). Затем эксперт аналогичным образом сформулировал диагностические правила для подзадачи «Осмотр» (5 признаков). После этого ему было предложено решить подзадачу с двумя признаками: предварительный диагноз по жалобам и предварительный диагноз по осмотру, значениями которых были наименования

классов, образовавшихся в результате решения подзадач «Жалобы» и «Осмотр», соответственно, а к наименованиям исходно названных и новых классов, добавлялись слова «Предварительный диагноз по жалобам и осмотру». Перечень таких классов составил значения одноименного признака в подзадаче непосредственно верхнего уровня, который был добавлен к исходно названным признакам группы «Анамнез» (6 исходных признаков), и так далее.

Декомпозиция по классам используется либо с самого начала, когда на этапе структуризации выясняется, что у некоторых признаков нет значений, допустимых для части классов, либо после выявления классифицирующих правил эксперта (возможно с использованием декомпозиции по признакам), если эксперт не смог однозначно разделить некоторую или всю совокупность объектов на исходно названные им классы, и отнес их сразу к нескольким классам. И в том, и в другом случае эксперт осуществляет классификацию объектов внутри каждой группы классов, и может использовать для этого как исходно названные им признаки (имеющие хотя бы одно допустимое значение для каждого класса такой группы), так и дополнительные признаки, необходимые для дифференциации объектов внутри каждой группы.

3. Выявление классифицирующих правил

Матрицы допустимости, построенные на этапе структуризации, позволяют определить для каждого объекта $a_i \in A$ кортеж его допустимости классам:

$$V_i = \left(\prod_{m=1}^M v^1(a_{im}), \prod_{m=1}^M v^2(a_{im}), \dots, \prod_{m=1}^M v^L(a_{im}) \right),$$

где $v^l(a_{im})$ означает допустимость $a_{im} \in Q_m$ классу C_l , и возможность проверить выполнимость для него одного из двух неявных правил:

$$\text{если } \forall l = \overline{1, L} \prod_{m=1}^M v^l(a_{im}) = 0, \text{ то } a_i \notin C_l, l = \overline{1, L}; \quad (1)$$

$$\text{если } \exists w \mid w = \overline{1, L} \prod_{m=1}^M v^w(a_{im}) \neq 0 \text{ и}$$

$$\forall l = \overline{1, L} \mid l \neq w \prod_{m=1}^M v^l(a_{im}) = 0, \text{ то } a_i \in C_w \quad (2)$$

Правило (1) означает, что объект a_i не принадлежит ни одному из заданных на данный момент классов и, следовательно, относится к классу C^{Un} , а правило (2) определяет единственный класс, к которому может принадлежать данный объект. Если для объекта выполняется одно из неявных правил, нет необходимости предъявлять его эксперту для классификации (если только структура задачи впоследствии не изменится). В противном случае, для такого объекта необходимо определить так называемые явные правила.

Явное правило для объекта выявляется посредством моделирования процесса принятия решений экспертом. Хотя значения всех признаков объекта известны, они сообщаются эксперту только по его запросу, кроме первого признака, выбранного случайным образом (что необходимо для последующей проверки согласованности правил). Очевидно, что такой информации недостаточно эксперту для классификации, и он либо запрашивает значение признака из множества Q , либо называет новый признак, значение которого хотел бы знать. В последнем случае эксперту предлагается выполнить процедуру явной структуризации, чтобы определить значения такого признака, их допустимость для классов, и, если применимо, упорядочить такие значения по их предпочтительности, в случае порядковой классификации, или по их типичности для классов, в случае номинально-порядковой классификации. Первое из названных значений показывается эксперту вместе с информацией, уже известной ему об объекте, и такой диалог продолжается, пока эксперт не примет одно из следующих решений:

а) выберет один или несколько (в случае номинально-порядковой или номинальной классификации) классов из C и/или назовет новый класс или классы (в вышеупомянутом случае);

б) укажет несовместимые значения признаков в описании объекта;

с) заявит, что не знает класс, к которому следует отнести данный объект.

Если в случае (а) эксперт называет новый класс или классы ему предлагается выполнить процедуру явной структуризации, чтобы определить допустимость значений всех признаков для каждого такого класса и упорядочить их по

предпочтительности, в случае порядковой классификации, или по их типичности для каждого такого класса, в случае номинально-порядковой классификации.

Обозначим через $Q_i^+ \in Q$ подмножество признаков, запрошенных экспертом при анализе объекта $a_i \in A$, а через $C_i^+ \in C$ – подмножество классов, к которым принадлежит любой объект с теми же значениями признаков из Q_i^+ , что и объект a_i .

Заметим, что, хотя эксперт игнорирует значения признаков из $Q \setminus Q_i^+$ (если $Q \setminus Q_i^+ \neq \emptyset$), разумно предположить, что он имеет в виду не просто «любые» значения этих признаков, а допустимые для каждого класса, к которому он отнес объект. Для непосредственно классифицированного объекта недопустимость конкретных значений проигнорированных признаков классам из C_i^+ отменяется, поскольку информация, полученная на этапе классификации в процессе анализа объекта, более надежна, чем на этапе структуризации, и для всех таких объектов может быть сформировано следующее общее правило:

$$\begin{aligned} &\text{если } \forall Q_m \in Q_i^+ \ a_{jm} = a_{im} \\ &\text{и } \forall Q_m \in Q_m \setminus Q_i^+, \forall C_l \in C_i^+ \ v^l(a_{jm}) \neq 0, \quad (3) \\ &\text{то } a_j \in C_i^+. \end{aligned}$$

Если решается задача порядковой классификации, это правило может быть распространено по доминированию (напомним, что, как отмечено выше в Примечании 4, в случае порядковой классификации эксперт может отнести непосредственно классифицируемый объект только к одному классу, который обозначим через C_x):

$$\begin{aligned} &\text{если } \forall Q_m \in Q_i^+ \ (a_{jm}, a_{im}) \in R_m \\ &\text{и } \forall Q_m \in Q_m \setminus Q_i^+, \forall y \leq x \ v^y(a_{jm}) \neq 0, \quad (4) \\ &\text{то } a_j \in C_y, 1 \leq y \leq x; \end{aligned}$$

$$\begin{aligned} &\text{если } \forall Q_m \in Q_i^+ \ (a_{im}, a_{jm}) \in R_m \\ &\text{и } \forall Q_m \in Q_m \setminus Q_i^+, \forall x \leq y \leq L \ v^y(a_{jm}) \neq 0, \quad (5) \\ &\text{то } a_j \in C_y, y = \overline{1, L}, x \leq y \leq L. \end{aligned}$$

В случае номинально-порядковой классификации, обозначим через $Q^{ord} \in Q$ подмножество

признаков, значения которых эксперт может упорядочить по типичности для классов, а подмножество признаков из Q_i^+ с упорядоченными таким образом значениями через $Q_i^{ord} \in Q^{ord} \cap Q_i^+$. Если $Q_i^{ord} \neq \emptyset$, правило (3) распространяется по доминированию с помощью двух следующих правил для каждого класса $C_l \in C_i^+$ и каждого класса $C_w \notin C_i^+$, соответственно:

$$\begin{aligned} &\text{если } \forall Q_m \in Q_i^+ \setminus Q_i^{ord} \ a_{jm} = a_{im} \\ &\text{и } \forall Q_m \in Q_i^{ord} (a_{jm}, a_{im}) \in R_m^l, \text{ то } a_j \in C_l; \end{aligned} \quad (6)$$

$$\begin{aligned} &\text{если } \forall Q_m \in Q_i^+ \setminus Q_i^{ord} \ a_{jm} = a_{im} \\ &\text{и } \forall Q_m \in Q_i^{ord} (a_{im}, a_{jm}) \in R_m^l, \text{ то } a_j \notin C_w. \end{aligned} \quad (7)$$

Правило (7) означает, что, поскольку эксперт отнес объект $a_i \in A$ к классам C_i^+ , это объект не принадлежит ни к какому иному классу, и это верно для любого другого объекта, значения которого не более типичны для таких иных классов, чем у объекта $a_i \in A$.

В случае (b) соответствующая несовместимая комбинация значений признаков относится к классу C^{Inc} . Таким образом, процедура классификации объектов заключается в следующем: во-первых, если $C^{Inc} \neq \emptyset$ для очередного неклассифицированного объекта проверяется, не включает ли его описание несовместимые комбинации значений признаков из C^{Inc} , и, если включает, объект считается классифицированным как «невозможный». Если не включает, проверяется, применимо ли к нему одно из неявных правил. Если применимо правило (1), объект относится к классу C^{Un} . Если применяется правило (2), объект относится к единственному допустимому классу. Если ни одно из явных правил неприменимо, проверяется применимость уже выявленных явных правил (3), а также правил (4) и (5), в случае задачи порядковой классификации, или правил (6) и (7), в случае номинально-порядковой классификации. Наконец, если ни одно из соответствующих явных правил неприменимо, явное правило для объекта выявляется описанным выше образом. Как только эксперт сформирует явное правило, проверяется его согласованность с каждым из ранее выявленных явных правил.

Эта процедура продолжается пока для каждого объекта, не содержащего несовместимых сочетания значений признаков, не будет существовать хотя бы одного неявного или явного правила. Такая полнота классификации требуется только с точностью до знаний эксперта. Это означает: если эксперт определяет, что объект не принадлежит ни к одному из классов из их существующего на тот момент множества, и не может назвать новый класс, такой объект считается классифицированным и принадлежащим классу C^{Un} .

4. Последствия изменения структуры задачи

Если эксперт называет в процессе классификации новый класс, он должен определить допустимость или недопустимость значений всех признаков для такого класса. Новый класс добавляется к C как C_L , где $L = L + 1$. Добавление нового класса не влияет на явные правила, выявленные ранее (если и пока не будет обнаружена их несогласованность с другими явными правилами), но может повлиять на неявные правила, которые были верны для объектов с меньшими номерами, чем номер последнего классифицированного объекта. Пусть один из таких объектов был ранее отнесен к классу C^{Un} в силу применимости к нему неявного правила (1). Если значения всех признаков у такого объекта допустимы для нового класса C_L , явное правило (1) становится для него неприменимым, а неявное правило (2) – применимым и, следовательно, он относится к классу C_L . Если какой-то объект классифицирован благодаря применимости к нему неявного правила (2), и значение какого-то признака у такого объекта недопустимо для нового класса C_L , применимость к нему неявного правила (2) сохраняется. Если же значения всех признаков у такого объекта допустимы для нового класса C_L , неявное правило (2) становится для него неприменимым. Проверка применимости неявных правил для каждого такого объекта проверяется автоматически, и, если неявные правила оказываются больше неприменимыми, проверяется применимость к нему уже выявленных явных правил. Если таких явных правил нет, объект предъявляется эксперту для классификации.

Если в процессе классификации объекта эксперт запрашивает новый признак, то он добавляется к Q как Q_M , где $M = M + 1$, и размер A увеличивается в n_M раз. Объекты увеличенного набора объектов A перенумеровываются в возрастающем лексикографическом порядке (см. Примечание 1). Добавление такого нового признака не влияет ни на явные правила, выявленные ранее, ни на неявные (1), если они были применимы. Однако если к какому-то объекту с меньшим номером, чем номер объекта, классифицированного последним, было применимо неявное правило (2), и новый признак имеет у такого объекта значение, недопустимое для того единственного класса, к которому он был отнесен, неявное правило (2) для такого объекта становится неприменимым, но неявное правило (1) становится применимым и относит его к классу C^{Un} .

Если эксперт отменяет недопустимость значения признака для какого-то класса, это может привести к неприменимости некоторых (или даже всех) неявных правил. В таком случае для каждого объекта с меньшим номером, чем номер объекта, классифицированного последним, автоматически проверяется применимость к нему неявных правил. Если для какого-то объекта ранее применимое к нему неявное правило стало неприменимым, проверяется применимость к нему уже выявленных явных правил. Если ни одно из явных правил неприменимо к такому объекту, эксперту предлагается его классифицировать.

Такие проверки выполняются после формирования каждого явного правила и до перехода к классификации следующего объекта.

5. Согласованность правил

Принадлежность каждого объекта к классу/классам может быть выведена из любого числа явных правил, но все они должны классифицировать его одинаково. Это условие означает согласованность правил классификации.

Пара явных правил (3) взаимно не согласована, если (а) пересечение подмножества признаков с конкретными значениями в части «если» первого правила и подмножества признаков с конкретными значениями в части «если» второго правила пусто; или (б) одни и те же признаки в части «если» обоих правил име-

ют одинаковые соответствующие значения и, в любом случае, содержат в части «то» разные классы, при условии, что значение каждого признака в части «если» первого правила допустимо для любого класса в части «то» второго правила. Значение каждого признака в части «если» второго правила допустимо для любого класса в части «то» первого правила. Шкала каждого признака, имеющего «любое значение» в обоих правилах имеет хотя бы одно значение, допустимое для каждого класса в части «то» первого и второго правила.

Рассмотрим два явных правила:

если $\forall Q_m \in Q_i^+ a_{jm} = a_{im}$ и

$\forall Q_m \in Q_m \setminus Q_i^+, \forall C_l \in C_i^+ v^l(a_{jm}) \neq 0$, то $a_j \in C_i^+$;

если $\forall Q_m \in Q_i^+ a_{jm} = a_{im}$

и $\forall Q_m \in Q_m \setminus Q_i^+, \forall C_l \in C_i^+ v^l(a_{jm}) \neq 0$, то $a_j \in C_i^+$;

$C_i^+ \neq C_i^+$.

Эти правила взаимно не согласованы, если или (а) $Q_i^+ \cap Q_i^+ = \emptyset$, или (б) $Q_i^+ \cap Q_i^+ \neq \emptyset$ и $\forall Q_m \in Q_i^+ \cap Q_i^+ a_{im} = a_{im}$ и в любом случае:

$\forall Q_m \in Q_i^+, \forall C_l \in C_i^+ v^l(a_{im}) \neq 0$,

$\forall Q_m \in Q_i^+, \forall C_l \in C_i^+ v^l(a_{im}) \neq 0$;

$\forall Q_m \in Q \setminus (Q_i^+ \cup Q_i^+), \forall C_l \in (C_i^+ \cup C_i^+) \exists q_{mk} | v_{mk}^l \neq 0$.

Если после выявления очередного явного правила обнаруживается его несогласованность с ранее выявленным правилом, эксперту показывается пара таких правил для анализа и исправления. Как только эксперт исправит какое-то явное правило, процедура проверки согласованности исправленного правила с другими правилами повторяется, пока все выявленные явные правила не будут согласованы. Если эксперт исправляет правило, выявленное раньше, среди классифицированных к данному моменту объектов осуществляется поиск объектов, которые теперь оказываются неклассифицированными. Эксперт должен классифицировать их до перехода к классификации еще не рассмотренных объектов.

В случае порядковой классификации проверяется также согласованность последнего явного правила с правилами (4) и (5), а в случае номинально-порядковой классификации — с правилами (6) и (7).

Заключение

Предложенный общий подход к решению задач многопризнаковой классификации на основе ВАР включает в себя достоинства методов, ранее разработанных на основе той же парадигмы, и свободен от присущих им недостатков. Он применим к задачам классификации разных типов – порядковой, номинально-порядковой и номинальной, а также к задачам большой размерности. Если задача классификации является порядковой или номинально-порядковой, для дополнительного сокращения числа объектов, непосредственно классифицируемых экспертом, можно использовать алгоритмы определения номера очередного объекта, подлежащего явной классификации экспертом, применяемые в методах CLASS и NORClass, соответственно.

В случае номинальной классификации эффективность метода STEPCLASS зависит, в основном, от двух факторов. Во-первых, это определенные экспертом допустимости/недопустимости значений признаков для классов. Так, например, при решении задачи классификации геологических пород [43] основное время было затрачено на ее структуризацию, в результате которой было названо 87 классов и 124 признака с числом значений от 5 до 10, а также на декомпозицию по признакам и по классам. При этом результирующая полная и согласованная классификация была получена на основе только неявных правил, и эксперту не пришлось непосредственно классифицировать ни один объект. Во-вторых, это число признаков, значения которых эксперт запрашивает при классификации каждого предъявленного ему объекта. Очевидно, что чем больше признаков, значения которых эксперт хочет узнать, чтобы классифицировать объект, тем больше объектов ему приходится непосредственно классифицировать. Однако опыт применения метода STEPCLASS показывает, что после непосредственной классификации нескольких первых объектов эксперт становится «более решительным» и запрашивает меньше признаков, значения которых ему необходимы для принятия решения.

Литература

1. Нейман фон Дж., Моргенштерн О. Теория игр и экономическое поведение / Пер. с англ. М.: Наука. 1970.
2. Кини Л. Р. Райфа Х. Принятие решений при многих критериях: предпочтения и замещения / Пер. с англ. Под ред. И.Ф. Шахнова. М.: Радио и связь. 1981.
3. Jacquet-Lagrange E., Siskos Y. Assessing a set of additive utility functions for multicriteria decision making: The UTA method. *EJOR* 10. 1982. P.151-164.
4. Roy B., Bouyssou D. Aide multicritère à la décision: méthodes et cas. Paris. Economica. 1993.
5. Mousseau V., Slowinski R., Zielniewicz P. A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research*. 2000. 27. P.757-777.
6. Figueira J., Mousseau V., Roy B.. Electre methods. *Multiple Criteria Decision Analysis: State of the Art Surveys*, eds. Figueira J., Greco S., M. Ehrgott. Boston: Springer-Verlag. 2005. P.133-162.
7. Greco S., Matarazzo B., Slowinski R.. Rough sets theory for multicriteria decision analysis. *EJOR*. 2001.129. P.1-47.
8. Greco S., Matarazzo B., Slowinski R. Dominance-Based Rough Set Approach as a Proper Way of Handling Graduality in Rough Set Theory. *T. Rough Sets* 7.2007. P.36-52.
9. Błaszczyński J., Greco S., Slowinski R. Inductive discovery of laws using monotonic rules. *Eng. Appl. of AI* 25(2). 2012. P.284-294.
10. Błaszczyński J., Greco S., Slowinski R. Multi-criteria classification – A new scheme for application of dominance-based decision rules. *EJOR* 181.2007.P. 1030–1044.
11. Zadeh L. Fuzzy sets. *Information Control*. 1965.8. P.338-353.
12. Fernandez E., Navarro J. A new approach to multi-criteria sorting based on fuzzy outranking relations: The THESEUS method. *EJOR*. 2011. 213(2). P.405-413.
13. Zopounidis C., Doumpos M. Multicriteria Classification and Sorting Methods: A Literature Review. *EJOR*. 2002. 138(2). P.229-246.
14. Perny P. Multicriteria filtering methods based on concordance and non-concordance principles. *Annals of Operations Research*. 1998. 80(1). P.137-166.
15. Belacel N. Multicriteria assignment method PROAFTN: Methodology and medical applications. *EJOR* 125. 2000. P.175-183.
16. Leger J., Martel J-M.. A multicriteria assignment procedure for a nominal sorting problematic. *EJOR*. 2002.125. P.175-183.
17. Belacel N., Boulassel M.R. Multicriteria Fuzzy Classification Procedure PROCFTN: Methodology and Medical Application, *Fuzzy Sets and Systems*.2004. 141(2). P.203-217.
18. Belacel N. The k-Closest Resemblance Approach for Multiple Criteria Classification problems, *Modelling, Computation and Optimization Information and Management Sciences*, eds. Thi Hoai, and Pham Dinh Tao. Hermes Sciences Publishing, London. 2004. P.525-534.
19. Rigopoulos G., Askounis D., Metaxiotis K. NeXClass: A Decision Support System for non-ordered Multicriteria Classification, *IJITDM*. 2010. 9(1). P.53-79.
20. Rigopoulos G., Anagnostopoulos K. Fuzzy Multicriteria Assignment for Nominal Classification Methodology and Application in Evaluation of Greek Bank's Electronic Payment Retailers, *IJITDM*. 2010. 9(3). P.1-18.
21. Yevseyeva I., Miettinen K., Salminen P., Lahdelma R. SMAA-Classification: A new method for nominal classification, Helsinki School of Economics, Working Paper. 2007. <http://hsepublib.lib.hse.fi/pdf/wp/w422.pdf>.
22. Costa A. S., Corrente S., Greco S., Figueira J. R., Borbinha J. A robust hierarchical nominal classification method based on similarity and dissimilarity. 2018.arXiv preprint arXiv:1812.08596.

23. Chen Ye, Multiple Criteria Decision Analysis: Classification Problems and Solutions. 2006. <http://etd.uwaterloo.ca/etd/y3chen2006.pdf>.
24. Ларичев О.И., Мечитов А.И., Мошкovich Е.М., Фуремс Е.М. Выявление экспертных знаний (процедуры и реализация). М.: Наука. 1989. 128 с.
25. Larichev O.I., Moshkovich H.M., Furems E.M., Mechitov A.I., Morgoev V.K. Knowledge Acquisition for the Construction of the Full and Contradiction Free Knowledge Bases. Groningen:Iec ProGAMMA. 1991. 230 p.
26. Furems Eugenia M. Domain Structuring For Knowledge-Based Multiattribute Classification (A Verbal Decision Analysis Approach). 2011. TOP 19(2). P.402-420.
27. Puppe F. 1998. Knowledge Reuse among Diagnostic Problem-Solving Methods in the Shell-Kit D3. International Journal of Human-Computer Studies.49. P.627-649.
28. Ларичев О.И., Мошкovich Е.М. Качественные методы принятия решений. М.: Физматлит. 1996.
29. Ashikhmin Ilya, Furems Eugenia, Petrovsky Alexey, Sternin Michael. Intelligent DSS under Verbal Decision Analysis//Encyclopedia of Decision Making and Decision Support Technologies (2 Volumes). Hershey • New York: Information Science Reference. 2008. P.514-527.
30. Larichev O.I., Moshkovich E.M., Furems E.M. Decision Support System 'Class'. New Direction in Research on Decision Making, eds. P. Lourenc and G. Svenson, North-Holland, Amsterdam. 1986. P. 305-315.
31. Larichev O.I., Moshkovich H.M.. An Approach to Ordinal Classification Problems, International Transactions of Operational Research. Vol 1,3.1994. P.375-385.
32. Larichev O., Asanov A., Naryzhny Y. Effectiveness evaluation of expert classification methods. EJOR. 2002. 138(2). P.260-273.
33. Ларичев О.И., Болотов А.А. Система ДИФКЛИАСС: построение полных и непротиворечивых баз экспертных знаний в задачах дифференциальной диагностики. Научно-техническая информация. Серия 2. 1996. С. 9-15.
34. Furems E., Gnedenko L. Stepclass—a system for eliciting expert knowledge and conducting expert analyses for solving diagnostic problems. Autom Doc Math Linguist 30(5). 1996. P.23–28. Translations of selected articles from Nauchno-Tekhnicheskaya Informatsiya.
35. Simon H. The New Science of Management Decision. Harper and Row. New York. 1960.
36. Фуремс Е.М. Многокритериальная порядковая классификация на основе метода STEPCLASS//Искусственный интеллект и принятие решений. 2012. № 4. С. 95-100.
37. Furems E.M. STEPCLASS-Based Approach to Multicriteria Sorting.//Scientific and Technical Information Processing. 2015. Vol 42. No 6. P. 87–96.
38. Фуремс Е.М. Модификация метода экспертной номинально-порядковой классификации// Искусственный интеллект и принятие решений. 2010. № 4. С.81-93.
39. Furems Eugenia M. Dominance-Based Extension Of STEPCLASS For Multiattribute Nominal Classification. IJITDM. 2013. Vol 12. No. 5. P.905–925.
40. Фуремс Е.М. Применение метода STEPCLASS к задачам классификации с иерархической структурой //Искусственный интеллект и принятие решений. 2014. № 4. С. 96-107.
41. Huang S. Dimensionality Reduction in Automatic Knowledge Acquisition: A Simple Greedy Search Approach. IEEE Transactions on Knowledge and Data Engineering. 2003.15(6). C.1364-1373.
42. Furems E., Sokolova L. Experts' knowledge acquisition for differential diagnostics of bronchial asthma in children in STEPCLASS environment. IJTPM. 2011. 11(1). С. 68-84.
43. Порожун В.И., Фарбирович В.П., Фуремс Е.М. Построение систем диагностики генетических типов терригенных и карбонатных пород в среде STEPCLASS//Научно-техническая информация. Серия 2. 1996. № 9. С. 27-31.

General Approach to Multiattribute Classification Problems Based on the Verbal Decision Analysis

Е. М. Furems

Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

Abstract. There are a number of methods for ordinal, nominal-ordinal and nominal classification problems within the Verbal Decision Analysis (VDA) paradigm. The VDA-based general approach to all such types of classification problems is proposed. This approach is implemented in the STEPCLASS method, which retains all advantages of above mentioned methods and eliminates their drawbacks. It includes, without limitations, such techniques as (1) initial domain structuring and defining the type of a classification problem; (2) flexible strategy of a dialogue with an expert in order to reveal the complete and consistent set of his/her classification rules with possibility of the problem domain structure extension; and (3) decomposition of large-size classification problem with subsequent aggregation of classification rules over the set of all attributes.

Keywords: verbal decision analysis, multiattribute classification, problem domain structuring, completeness and consistency of classification rules, classification problems with hierarchical structure and/or large-size classification problems.

DOI 10.14357/20718594190406

References

1. von Neumann J. and Morgenstern O. Theory of games and economic behavior. – 2nd ed. – Princeton: Princeton University Press, 1947.
2. R. Keeney R.L., Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. – N.Y.: J. Wiley, 1976.
3. Jacquet-Lagrange E., Siskos Y. 1982. Assessing a set of additive utility functions for multicriteria decision making: The UTA method, *EJOR* 10:151-164.
4. Roy B., Bouyssou D., 1993. Aide multicritère à la décision: méthodes et cas. Paris: Economica.
5. Mousseau V., Slowinski R., Zielniewicz P. 2000. A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research* 27:757-777.
6. Figueira J., Mousseau V., Roy B. 2005. Electre methods. Multiple Criteria Decision Analysis: State of the Art Surveys, eds. Figueira J., Greco S., M. Ehrgott. Boston: Springer-Verlag, 133-162 p.
7. Greco S., Matarazzo B., Slowinski R. 2001. Rough sets theory for multicriteria decision analysis. *EJOR*. 129:1-47.
8. Greco S., Matarazzo B., Slowinski R. 2007. Dominance-Based Rough Set Approach as a Proper Way of Handling Graduality in Rough Set Theory. *T. Rough Sets* 7(2007) 36-52.
9. Błaszczyński J., Greco S., Slowinski R. 2012. Inductive discovery of laws using monotonic rules. *Eng. Appl. of AI* 25(2): 284-294.
10. Błaszczyński J., Greco S., Slowinski R. 2007. Multicriteria classification – A new scheme for application of dominance-based decision rules, *EJOR* 181: 1030–1044
11. Zadeh L. 1965. Fuzzy sets. *Information Control*, 8: 338-353
12. Fernandez E., Navarro J. 2011. A new approach to multicriteria sorting based on fuzzy outranking relations: The THESEUS method. *EJOR*, 213(2): 405-413
13. Zopounidis C., Doumpos M. 2002. Multicriteria Classification and Sorting Methods: A Literature Review. *EJOR*. 138(2): 229-246.
14. Perny P. 1998. Multicriteria filtering methods based on concordance and non-concordance principles. *Annals of Operations Research* 80(1): 137-166.
15. Belacel N. 2000. Multicriteria assignment method PROFTN: Methodology and medical applications, *EJOR* 125: 175-183.
16. Leger J., Martel J-M. 2002. A multicriteria assignment procedure for a nominal sorting problematic. *EJOR* 125: 175-183.
17. Belacel N., Boulassel M.R. 2004. Multicriteria Fuzzy Classification Procedure PROCFTN: Methodology and Medical Application, *Fuzzy Sets and Systems*, 141(2): 203-217.
18. Belacel N. 2004. The k-Closest Resemblance Approach for Multiple Criteria Classification problems, *Modelling, Computation and Optimization Information and Management Sciences*, eds. Thi Hoai, and Pham Dinh Tao. Hermes Sciences Publishing, London, 525-534.
19. Rigopoulos G., Askounis D., Metaxiotis K. 2010. NeX-CLASS: A Decision Support System for non-ordered Multicriteria Classification, *IJITDM*, 9(1): 53-79.
20. Rigopoulos G., Anagnostopoulos K. 2010. Fuzzy Multicriteria Assignment for Nominal Classification Methodology and Application in Evaluation of Greek Bank's Electronic Payment Retailers, *IJITDM*, 9(3):1-18.
21. Yevseyeva I., Miettinen K., Salminen P., Lahdelma R. 2007. SMAA-Classification: A new method for nominal classification, Helsinki School of Economics, Working Paper, <http://hsepublib.lib.hse.fi/pdf/wp/w422.pdf>.
22. Costa A. S., Corrente S., Greco S., Figueira J. R., Borbinha J. 2018. A robust hierarchical nominal classification method based on similarity and dissimilarity. *arXiv preprint arXiv:1812.08596*.
23. Chen Ye, Multiple Criteria Decision Analysis: Classification Problems and Solutions, 2006, <http://etd.uwaterloo.ca/etd/y3chen2006.pdf>.
24. Larichev O.I., Mechtov A.I., Moshkovich E.M., Furems E.M. 1989. Vyavlenie ekspertnykh znaniy (protsedury i realizatsiya) [Expert's knowledge acquisition (procedures and implementation)]. – M.: Nauka, 128 P.
25. Larichev O.I., Moshkovich H.M., Furems E.M., Mechtov A.I., Morgoev V.K. 1991. Knowledge Acquisition for the Construction of the Full and Contradiction Free Knowledge Bases. – Groningen: lec ProGAMMA, 230 p.
26. Furems Eugenia M. 2011. Domain Structuring For Knowledge-Based Multiattribute Classification (A Verbal Decision Analysis Approach). *TOP* 19(2): 402-420.
27. Puppe F. 1998. Knowledge Reuse among Diagnostic Problem-Solving Methods in the Shell-Kit D3. *International Journal of Human-Computer Studies*, 49: 627-649.
28. Larichev, O.I. and Moshkovich, E.M. 1996. Kachestvennye metody prinyatiya reshenii (Qualitative Methods of Decision-Making), Moscow: Fizmatlit.
29. Ashikhmin Ilya, Furems Eugenia, Petrovsky Alexey, Sterin Michael. 2008. Intelligent DSS under Verbal Decision Analysis/ //Encyclopedia of Decision Making and Decision Support Technologies (2 Volumes). Hershey • New York: Information Science Reference. 514-527.
30. Larichev O.I., Moshkovich E.M., Furems E.M. 1986. Decision Support System 'Class'. New Direction in Research on Decision Making, eds. P. Lourenc and G. Svenson, North-Holland, Amsterdam, 305-315.
31. Larichev O.I., Moshkovich H.M. 1994. An Approach to Ordinal Classification Problems, *International Transactions of Operational Research*, Vol. 1, 3: 375-385.
32. Larichev O., Asanov A., Naryzhny Y. 2002. Effectiveness evaluation of expert classification methods, *EJOR*, 138(2): 260-273.
33. Larichev, O.I. and Bolotov, A.A. 1996. Systema DIFKCLASS: postroenie polnykh i neprotivorechivyykh baz ekspertnykh znaniy v zadachah differentsialnoy diagnostiki [The DIFKCLASS system: Construction of Complete and Non-contradictory expertise databases in problems of differential diagnostics], *Nauchn. Tekhn. Inform.*, Ser. 2. vol.9: 9–15.
34. Furems E., Gnedenko L. 1996. Stepclass—a system for eliciting expert knowledge and conducting expert analyses for solving diagnostic problems. *Autom Doc Math Linguist* 30(5):23–28. Translations of selected articles from Nauchno-Tekhnicheskaya Informatsiya
35. Simon H.. 1960. The New Science of Management Decision. Harper and Row, New York.

36. Furems E.M, 2012. Mnogokriterialnaya poryadkovaya klassifikatsiya na osnove metoda STEPCLASS (Multicriteria ordinal classification on the basis of the STEPCLASS method. *Iskusstvennyi intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making], 4: 95-100.
37. Furems E.M. 2015. STEPCLASS-Based Approach to Multicriteria Sorting.//*Scientific and Technical Information Processing*, vol. 42, No. 6: 87–96.
38. Furems E.M. 2010. Modifikatsiya metoda erspertnoy nominalno-poryadkovoy klassifikatsii [Modification of the method for nominal-ordering classification], *Iskusstvennyi intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making], 4: 81-93.
39. Furems Eugenia M. 2013. Dominance-Based Extension Of STEPCLASS For Multiattribute Nominal Classification. *IJITDM*, vol. 12, No. 5: 905–925.
40. Furems E.M.. 2014. Primenenie metoda STEPCLASS k zadacham klassifikatsii s ierarhicheskoy strukturoy [Application of the STEPCLASS method for classification problem with hierarchical structure]. *Iskusstvennyi intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making], 4: 96-107.
41. Huang S. 2003. Dimensionality Reduction in Automatic Knowledge Acquisition: A Simple Greedy Search Approach. *IEEE Transactions on Knowledge and Data Engineering*, 15(6): 1364-1373.
42. Furems E., Sokolova L. 2011. Experts' knowledge acquisition for differential diagnostics of bronchial asthma in children in STEPCLASS environment. *IJTPM*, 11(1): 68-84.
43. Poroskun V.I., Farbirovich V.P., Furems E.M. 1996. Postroenie system diagnostiki geneticheskikh types terrigennykh i karbonatnykh porod v srede STEPCLASS [Development of the system for diagnostics of genetic types of terrigenous and carbonate deposits in the STEPCLASS environment]. *Nauchn. Tekhn. Inform., Ser. 2*. vol. 9: 27-31.