ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

УДК 007:681.512.2

КОРРЕКЦИЯ КОВАРИАНТНОГО ДРЕЙФА КОНЦЕПЦИИ ДЛЯ АНСАМБЛЕЙ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

И. Ю. Каширин, д.т.н., профессор кафедры ВПМ РГРТУ, Рязань, Россия;
orcid.org/0000-0003-1694-7410, e-mail: igor-kashirin@mail.ru

Рассмотрен новый подход к обнаружению и коррекции одной из разновидностей дрейфа данных в моделях машинного обучения, а именно ковариантного дрейфа концепции. Подход предполагает, что модель машинного обучения спроектирована как ансамбль моделей различных уровней. Методом сбора ансамбля является композиционный беггинг-метод.

Беггинг-метод использует в качестве компонентов ансамбля вначале слабые однотипные модели, затем применяется ряд итераций, позволяющих повышать точность результирующей модели до некоторого уровня, приемлемого для решения задачи прогнозирования по точности и вычислительной сложности.

Исследуются различные формулы дрейфа концепции, основанные на условной и безусловной вероятностях получения целевой переменной в зависимости от данных вектора признаков во входном наборе данных. Вводятся понятия положительного и негативного дрейфа концепции в зависимости от принадлежности к соответствующему классу, используемому в прогнозировании.

Новый подход использует понятийную базу знаний предметной области, позволяющую априори классифицировать элементы вектора признаков в форме родовидовой таксономии. Классифицированный вектор признаков представляет собой иерархическую структуру, позволяющую с помощью алгоритма бутстрэпа формировать подвыборки признаков (фолды) для предварительного обучения слабых моделей первого и второго уровней. Фолды в этом случае можно упорядочить и затем использовать полученный порядок для выявления и компенсации ковариантного дрейфа концепции рабочей модели.

В качестве примера для экспериментальных исследований взята предметная область услуг авиаперевозок из международного репозитория Kaggle. Программная реализация выполнена с применением инструментария Spider v.4 на языке Python v.4. Результаты проведенных экспериментов показывают эффективность нового подхода к коррекции дрейфа концепции.

***Целью работы** является получение нового подхода к идентификации и исправлению ковариантного дрейфа концепции, дающего возможность коррекции дрейфа в ансамблях моделей машинного обучения.*

***Ключевые слова:** дрейф концепции, ансамбли моделей машинного обучения, большие данные, точность прогнозирования, база знаний, онтологические модели знаний.*

DOI: 10.21667/1995-4565-2023-85-44-53

Введение

Проблема дрейфа данных и концепций в моделях машинного обучения возникает с течением времени как ухудшение точности прогнозирования или классификации [1, 2]. В этом случае обученная на более ранних данных модель не обеспечивает своих функций на более поздних входных наборах данных.

Перечислим причины изменения соотношений в текущих данных в сравнении с прежними данными:

– изменение источников входного набора данных;
– демографические изменения с течением лет;
– появление новых высоких технологий;
– выявление ранее неизученных факторов влияния в анализируемой предметной области;
– влияние периодических природных явлений;
– перевод фокуса интересов пользователей на новые факторы.

Для преодоления дрейфа концепций в анализируемой предметной области необходимо тщательно реализовать все этапы жизненного цикла разработки и эксплуатации моделей машинного обучения. В наиболее известном подходе проектирования средств Data Mining (CRISP-SM [3]) жизненный цикл планировался как последовательность таких работ, как понимание предметной области, понимание данных, подготовка данных, создание модели, оценка и внедрение.

Теоретическая часть

С выявлением серьезности проблемы дрейфа концепции жизненный цикл имеет смысл представлять новой последовательностью (рисунок 1). На рисунке 1 сплошной линией обозначена последовательность этапов проектирования и эксплуатации моделей, адаптированных к дрейфу концепции. Пунктирная линия соответствует применению автоматизированных средств мониторинга, переоценки и online-исправления моделей. Из рисунка следует, что мониторинг данных требуется на каждом из этапов проектирования и эксплуатации прогнозирующих моделей.

В настоящей статье публикуется новый подход к идентификации проблемы дрейфа концепции, ориентированный на ансамбли моделей машинного обучения, собранные с помощью беггинг-метода [4]. Подход назван методом семантической фильтрации (МСФ). Для изложения сути подхода будет применяться формальный аппарат, использующий понятие распределения вероятностей [5].

МСФ используется в задачах бинарной классификации, где два подкласса целевых меток при прогнозировании интерпретируются как «положительный результат» и «отрицательный результат», соответствующие желательному развитию событий и нежелательному развитию событий с точки зрения лица, принимающего решение.

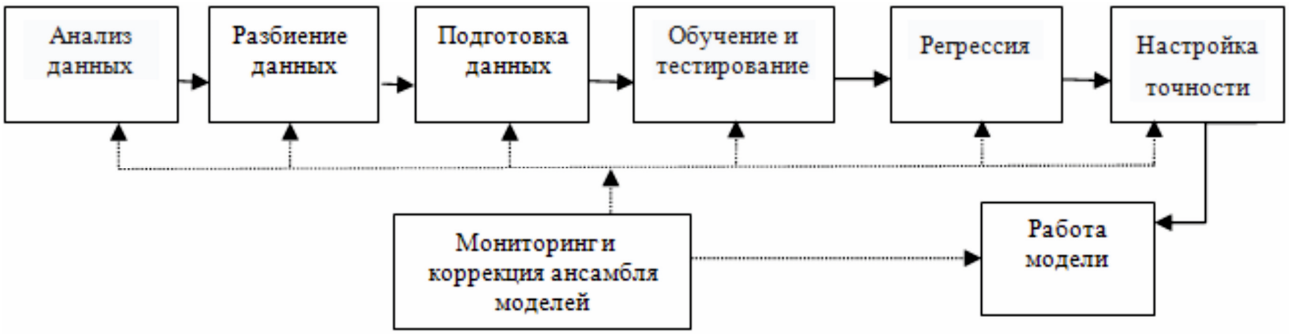

Рисунок 1 – Схема разработки и эксплуатации моделей машинного обучения
Figure 1 – Scheme of machine learning models development and operation

То же самое можно отнести и к векторам входных признаков, изменение которых при ковариантном дрейфе можно разделить на «положительный сдвиг», «нейтральный сдвиг» и «отрицательный сдвиг».

Далее приводится формальное описание основных понятий, на которых базируется МСФ, и кратко перечислены существующие методы обнаружения дрейфа концепции.

Пусть $P(X)$ – безусловная вероятность получения во входном наборе данных вектора признаков X, а $P(y|X)$ – условная вероятность получения значений целевой переменной y при наличии входных признаков X.

При исследовании дрейфа данных относительно прогнозирующей или классифицирующей ML-модели выделяют два основных типа дрейфа: *дрейф экземпляров признаков* и *дрейф системы понятий*.

Дрейф экземпляров признаков (дрейф данных) предполагает изменение $P(X)$, вызванное изменением входного набора данных. Зависимость целевой переменной y от вектора признаков X остается неизменной. При этом изменение $P(X)$ может повлечь за собой изменение $P(y|X)$. Для идентификации дрейфа данных вначале необходимо выявить изменение в распределении данных вектора X.

Например, дрейфом экземпляров признаков можно считать изменение значений признака «количественное снижение месячных закупок» или значений признака «количество посещений сайта за последние 24 часа».

Дрейф системы понятий (дрейф концепции) – это изменение $P(y|X)$, т.е. изменение зависимости целевой переменной y от вектора признаков X. Дрейф системы понятий предполагает, что само распределение признаков может быть неизменным, однако закономерности, выявляемые ML-моделью, начинают выявляться менее достоверно. В результате обученные модели становятся неэффективными. Первоначальное исследование дрейфа концепции предполагает идентификацию изменений в распределении целевой переменной y и/или изменение зависимости y от X.

Базовое соотношение, идентифицирующее дрейф системы понятий, определяется как изменение распределения X и y относительно изначального момента времени t по прошествии достаточно длительного времени Δt:

$$P_t(X, y) \neq P_{t+\Delta t}(X, y).$$

Примером дрейфа системы понятий является ввод санкций США против независимых стран. Введение таких санкций привело к укреплению популярности юаня (рост объема продаж) как международной валюты, а страны Европы – к существенной потере договороспособности (снижение числа договоров) и популярности, а также частичной потере спроса на их продукцию (снижение дохода).

Для анализа изменений входных наборов данных требуется упорядоченная во времени последовательность их экземпляров. Различают дрейф в потоковых данных и дрейф в пакетных (накопленных) данных.

Дрейф концепции обычно рассматривается в контексте потокового обучения. Входной поток данных определяется как непрерывная, потенциально неограниченная последовательность данных с соответствующей временной шкалой. Это отличается от дрейфа набора данных в контексте пакетного обучения, когда данные полностью присутствуют в памяти и обрабатываются сразу.

Для исследования разновидностей дрейфа системы понятий используются следующие обозначения:

y, y^+, y^- − целевые метки, соответственно: положительные вместе с отрицательными, положительные, отрицательные;

$P_t(X)$ − распределение вероятностей входных данных сразу после обучения ML-модели;

$P_t(y)$ − предварительное распределение вероятностей целевых меток;

$P_t(y|X)$ − апостериорное распределение вероятностей целевых меток;

$P_t(X|y)$ − распределение плотности вероятности в зависимости от класса.

Здесь целевые метки y^+ − это истинно позитивные прогнозы (TP) вместе с ошибочно позитивными прогнозами (FP), а y^- − это истинно отрицательные прогнозы (TN) совместно с ошибочно отрицательными прогнозами (FN) из матрицы ошибок (рисунок 2).

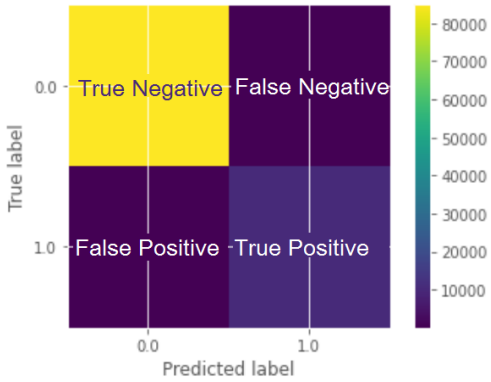

Рисунок 2 – Матрица ошибок
Figure 2 – Errors matrix

Если множество функций дрейфа S определить на двух аргументах P_t и $P_{t+\Delta t}$, то функции дрейфа концепции S_i $(P^i_t, P^i_{t+\Delta t})$ могут в своем разнообразии ограничиваться свойствами, представленными в таблице 1. Для удобства чтения верхние индексы i распределений вероятности в таблице опущены.

Таблица 1 – Формулы дрейфа концепции
Table 1 – Concept drift formulas

Ф-ция	**Формула дрейфа концепции**	**Наименование дрейфа концепции**
S_1	$P_t(y\|X) = P_{t+\Delta t}(y\|X)$	Отсутствие дрейфа
S_2	$P_t(y^-\|X) \neq P_{t+\Delta t}(y^-\|X)$	Действительный негативный дрейф
S_3	$P_t(y^+\|X) \neq P_{t+\Delta t}(y^+\|X)$	Действительный положительный дрейф
S_4	$P_t(X) \neq P_{t+\Delta t}(X)$	Ковариантный дрейф
S_5	$P_t(y) \neq P_{t+\Delta t}(y)$	Априорный дрейф
S_6	$P_t(X) = P_{t+\Delta t}(X)$	Отсутствие ковариантного дрейфа
S_7	$\exists x \in X, P_t(y\|x) \in C_1 \,\&\, P_{t+\Delta t}(y\|x) \in C_2$	Фрагментарный дрейф из класса C_1 в класс C_2
S_8	$\exists x \in X, P_t(y\|x) \in C_2 \,\&\, P_{t+\Delta t}(y\|x) \in C_1$	Фрагментарный дрейф из класса C_2 в класс C_1
S_9	$\exists x \in X, P_t(y\|x) \neq P_{t+\Delta t}(y\|x)$	Фрагментарный дрейф в любом направлении
S_{10}	$\forall x \in X, P_t(y\|x) \in C_1 \,\&\, P_{t+\Delta t}(y\|x) \in C_2$	Полный дрейф из класса C_1 в класс C_2
S_{11}	$\forall x \in X, P_t(y\|x) \in C_2 \,\&\, P_{t+\Delta t}(y\|x) \in C_1$	Полный дрейф из класса C_2 в класс C_1
S_{12}	$\exists x \in X, P_t(y\|x) \in C_1 \,\&\, P_{t+\Delta t}(y\|x) \in C_2$ $\&\exists z \in X, P_t(y\|z) \in C_2 \,\&\, P_{t+\Delta t}(y\|z) \in C_1$	Фрагментарный дрейф в разных классах
S_{13}	$P_t(X_1) \neq P_{t+\Delta t}(X_1) \,\&\, P_t(X_2) = P_{t+\Delta t}(X_2)$	Виртуальный дрейф в разных классах
S_{14}	$(C_t = C_1 \cup C_2 \,\&\, C_{t+\Delta t}) = (C_1 \cup C_2 \cup C_3)$	Концептуальная эволюция

На множестве приведенных функций дрейфа можно дать следующие определения видам концептуального дрейфа $S_a - S_m$:

$S_a = S_1$, $S_b = S_2 \,\&\, S_4 \,\&\, S_5$, $S_c = S_3 \,\&\, S_4 \,\&\, S_5$,

$S_S = S_2 \,\&\, S_3 \,\&\, S_4 \,\&\, S_5$, $S_e = S_5 \,\&\, S_6$, $S_f = S_7 \,\&\, S_8$,

$S_g = S_{10} \,\&\, S_{11}$, $S_h = S_{12}$ $S_i = S_{13}$, $S_k = S_1 \,\&\, S_4$, $S_m = S_{14}$.

Далее (рисунок 3) приведено графическое представление различных видов концептуального дрейфа $S_a - S_m$.

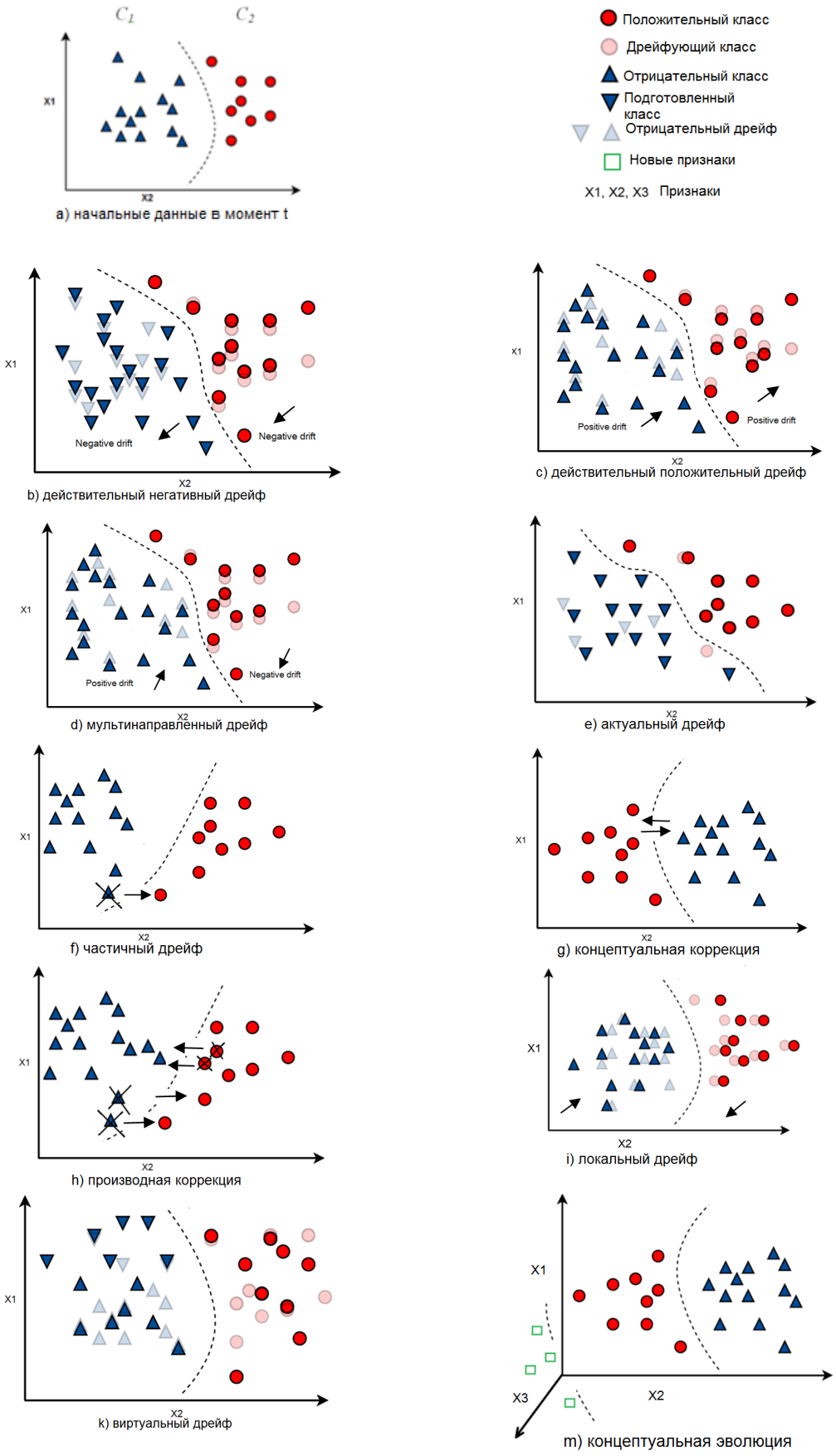

Рисунок 3 – Схемы дрейфа концепции
Figure 3 – Concept drift diagrams

Подход концептуальной фрагментации, предлагаемый здесь автором, заключается в использовании моделей знаний [6-8], в частности онтологических моделей, для разбиения входных наборов данных на локальные фрагменты.

Этот подход более эффективен при разработке и исследовании беггинг-ансамблей моделей машинного обучения. Беггинг-метод использует в качестве компонентов ансамбля вначале слабые однотипные модели, затем применяется ряд итераций, позволяющих повышать точность результирующей модели до некоторого уровня, приемлемого для решения задачи прогнозирования по точности и вычислительной сложности. При проектировании первоначальных (слабых) моделей исходный набор данных разбивается на локализованные фрагменты, которые включают лишь некоторые признаки вектора признаков X и некоторые выборки записей входного набора данных. Чаще всего предполагается случайная выборка таких фрагментов.

В предлагаемом подходе используется строгая упорядоченность разделения входных данных на выборки (фолды), основанная на последовательном обходе концептуальных подмножеств признаков вектора признаков X.

Признаки входного набора данных являются оцифрованными свойствами концептов или отношений модели знаний предметной области, в которой стоит задача классификации или прогнозирования целевых значений. Таким образом, концепты или отношения выделяют из всего множества признаков локальные подмножества, характеризующие эти концепты или отношения. Если использовать родовидовую и причинно-следственную таксономии [9], локальные подмножества признаков будут вложенными друг в друга подмножествами. Чем более общим по своей семантике является концепт, тем большее число элементов входит в соответствующее этому концепту множество признаков.

Если разделить входной датафрейм на кластеры, то с учетом таксономии концептов предметной области это можно сделать, классифицируя данные по конкретизации концептов снизу вверх или сверху вниз. Может получиться так, что в дрейфе находится часть признаков одного концепта и часть признаков другого концепта, причем эти два концепта не имеют общего предка. Это может свидетельствовать о следующих ситуациях.

1. Некорректна структура таксономии.
2. Нет достаточного количества признаков для точной родовидовой классификации.
3. Таксономия недостаточно подробно классифицирована.
4. Существует недоопределенный концепт по родовидовой иерархии смежного наследования. В литературе иногда огрубленно определяют смежное наследование как множественное наследование или ромбовидное наследование.
5. Рамки времени определены для двух признаков недостаточно точно. Т.е. в реальности для каждого из них время наступления, продолжения и окончания дрейфа различно.
6. Два признака совершенно разных концептов случайно оказались в совместном дрейфе в одних и тех же рамках времени.

Все перечисленные ситуации в большинстве случаев требуют дополнительной работы с моделью данных предметной области. Предметная область описывается вначале в форме какой-нибудь эффективной модели знаний, например онтологической [10-12], с применением языковых средств OWL. Затем эта модель знаний подвергается фрагментации, предполагающей выделение относительно самостоятельных фрагментов знаний, соответствующих семантической интерпретации данных входного обучающего набора [13-14]. Самостоятельные (локальные) фрагменты модели знаний связаны с данными входного обучающего набора отношением «концепт - атрибут». Отдельные концепты, обладая соответствующим списком атрибутов, представляют собой относительно самостоятельный семантический фрагмент. Если модель знаний предметной области спроектирована правильно, все множество фрагментов полностью покрывает весь список признаков входного обучающего набора данных модели машинного обучения. Теперь исследовать признаки на наличие дрейфа данных можно без использования полного их перебора, а фрагментами сразу по нескольку признаков,

объединенных единой парадигмой. Это, в то же время, не исключает последовательного уточнения роли каждого из признаков семантического фрагмента в выявленном дрейфе данных.

В качестве примера была исследована предметная область авиаперевозок (рисунок 4). В ней главным концептом является понятие «Полет», которым посредством концептуальных отношений могут быть «Потребитель» авиауслуг и «Сервис», которые, в конце концов, связываются с конкретными людьми и известными характеристиками.

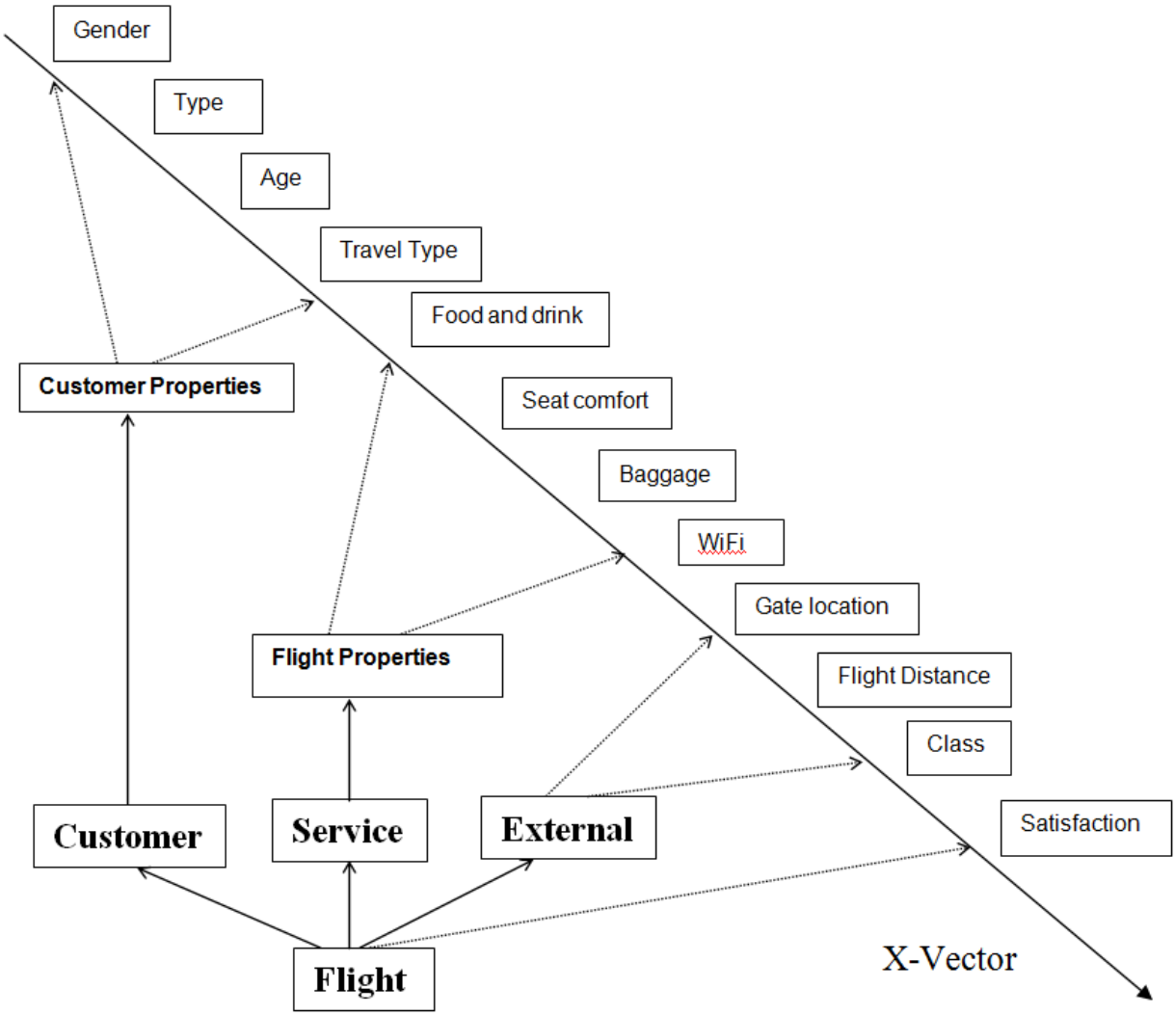

Рисунок 4 – Структуризация X-вектора для предметной области авиаперевозок
Figure 4 – Structuring of X-vector for the subject area of air transportation

На рисунке 4 отображены не только главные концепты сферы авиауслуг, но и отдельные признаки, которые могут быть представлены в числовом выражении: «Пол», «Тип пользователя», «Возраст», …, «Дистанция перелета», «Экономический класс», «Удовлетворенность авиауслугами». Теперь все числовые признаки (таблица 2) можно фрагментировать согласно структуризации.

На рисунке 4 приведен X-вектор признаков, состоящий из множества {«Gender», «Type», «Age», …, «Class», «Satisfaction»}. В классическом беггинг-методе выбор признаков этого вектора для проектирования первого уровня ансамбля моделей является случайным. Такой выбор является вполне эффективным фактором для построения ансамбля моделей, но становится препятствием при наличии дрейфа концепции. В этом случае заново собрать ансамбль моделей по той же самой схеме весьма затруднительно, тем более, что со временем эксплуатации ансамбля данные входного набора существенно изменяются.

Таблица 2 – Фрагмент данных для обучения
Table 2 – Data fragment for training

Gender	Customer Type	Age	Type of Travel	Class	Flight Distance	Inflight wifi service	Gate location	Food and drink	Seat comfort	On-board service	Baggage handling	satisfaction
Female	Loyal	52	Business	Eco	160	5	4	3	3	5	5	yes
Female	Loyal	36	Rest	Business	2863	1	1	5	5	4	4	yes
Male	Disloyal	20	Rest	Eco	192	2	4	2	2	4	3	no
Male	Loyal	44	Business	Business	3377	0	2	3	4	1	1	yes
Male	Disloyal	34	Business	Business	526	3	1	4	4	3	4	no
Male	Loyal	14	Business	Business	1127	3	3	4	4	3	5	yes

Предложенный в настоящей статье подход позволяет компенсировать рассмотренный недостаток для дрейфа концепций, определяемого формулами S_7, S_8, S_9, S_{12} таблицы 1. Для этого на этапе разбиения данных (рисунок 1) используется следующая онтологическая модель данных, представляющая собой родовидовую таксономию.

```
<owl:Class rdf:ID="Customer">
  <rdfs:subClassOf rdf:resource="#Flight" />
</owl:Class>
<owl:Class rdf:ID="Service">
  <rdfs:subClassOf rdf:resource="#Flight" />
</owl:Class>
<owl:Class rdf:ID="External">
  <rdfs:subClassOf rdf:resource="#Flight" />
</owl:Class>
<owl:Class rdf:ID="Flight Properties">
  <rdfs:subClassOf rdf:resource="#Service" />
</owl:Class>
<owl:Class rdf:ID="Customer Properties">
  <rdfs:subClassOf rdf:resource="#Customer" />
</owl:Class>
```

Согласно приведенной таксономии изначальный выбор признаков для X-вектора при беггинг-методе сборки ансамбля моделей осуществляется по правилу обхода признаков по порядку подклассов таксономии, т.е., например, сначала строится слабая модель для признаков класса «Customer», затем для класса «Flight Properties», и в конце для класса «External». Порядок и уровень конкретизации классов могут быть выбраны дата-сайентистом различными способами в зависимости от особенностей предметной области применения ансамбля моделей. Однако в любом случае этот порядок необходимо сохранить для последующего использования в случае обнаружения дрейфа концепции по схемам S_7, S_8, S_9, S_{12} таблицы 1. Коррекция дрейфа концепции производится в той же последовательности, в которой собирался изначальный ансамбль моделей.

Экспериментальные исследования

Как показали практические исследования, максимальный дрейф концепции (ROC-AUC: 0,882 → 0,80 за год) был идентифицирован для подмножества признаков owl-подкласса «Customer». При обновлении этих данных в исходном наборе данных характеристику точности ROC-AUC с использованием предложенного подхода удалось существенно компенсировать (ROC-AUC: 0,882 → 0,84), что позволяет сделать вывод об эффективности предложенного подхода и его прикладном значении.

Библиографический список

1. **Asghari M., Sierra-Sosa S., Telahun M., Kumar A., Elmaghraby A.S.** Aggregate density-based concept drift identification for dynamic sensor Data models // Neural Comput. Appl. 33 (8) (2021), pp. 3267-3279.

2. **Каширин И. Ю.** Data Mining с использованием иерархических чисел в ретроспективной диагностике // Вестник Рязанского государственного радиотехнического университета. 2022. № 79. С.118-126.

3. **Schröerab Ch., Kruseb F., Gómezb J. M.** A Systematic Literature Review on Applying CRISP-SM Process Model // ProceSia Computer Science 181 (2021), pp. 526-534.

4. **Breiman L.** Bagging predictors. Machine Learning, 24:123–140, 1996.

5. **Tsiakmaki M., Kostopoulos G., Kotsiantis S. and Ragos O.** Transfer learning from deep neural networks for predicting student performance // Applied Sciences, vol. 10, no. 6 (2020) pp. 2145–2218.

6. **Bayram F., Ahmed B. S., Kassler A.** From concept drift to model degradation: An overview on performance-aware drift detectors // Knowledge-Based Systems. 245 (2022) 108632

7. **Ali T. M., Nawaz A., Rehman A. et al.** A sequential machine learning-cum-attentions mechanism for effective segmentation of brain tumor *// Frontiers in Oncology*, vol. 12, pp. 1-10, 2022.

8. **Tsai F. S., Cabrilo S., Chou H. H., Hu F., Tang A. S.**. Open innovation and SME performance: the roles of reverse knowledge sharing and stakeholder relationships // Journal of Business Research. 2022, vol. 148, pp. 433-443.

9. **Xia K., Kai-Zhan Lee, Bengio Y., Bareinboim E.** The causal-neural connection: Expressiveness, learnability, and inference // Advances in Neural Information Processing Systems, 34:10823-10836, 2021.

10. **Wang T. S., Parsia B., Hendler J.** A Survey of the Web Ontology Landscape // The Semantic Web – ISWC 2006. Lecture Notes in Computer Science, vol. 4273, 682 p.

11. **Kashirin I. Yu., Filatov I. Yu**. FormalizeS Sescription Of Intuitive Perception Of Spatial Situations //2019 8th MeSiterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 2019, pp. 1-4.

12. **Каширин И. Ю.** Нейронные сети, использующие модели знаний // Вестник Рязанского государственного радиотехнического университета. 2021. № 75. С.71-84.

13. **Сатыбалдина Д. Ж., Овечкин Г. В., Калымова К. А.** Система распознавания статических жестов рук с использованием камеры глубины // Вестник рязанского государственного радиотехнического университета. 2020. № 72. С. 93-105.

14. **Phuong T. Nguyen, Davide Di Ruscio, Alfonso Pierantonio, Juri Di Rocco, and Ludovico Iovino.** Convolutional neural networks for enhanced classification mechanisms of metamodels // Journal of Systems and Software 172 (2021).

USC 007:681.512.2

CORRECTION OF COVARIANT DRIFT OF THE CONCEPT FOR ENSEMBLES OF MACHINE LEARNING MODELS

I. Yu. Kashirin, Dr. Sc. (Tech.), full professor, RSREU, Ryazan, Russia;
orciS.org/0000-0003-1694-7410, e-mail: igor-kashirin@mail.ru

The article discusses a new approach to detect and correct one of data drift types in machine learning models, namely covariant conceptual drift. The approach assumes that machine learning model is designed as a set of models of various levels. The method of collecting the ensemble is a compositional begging method.

The Begging method first uses weak models of the same type as ensemble components, then a number of iterations are used to increase the accuracy of resulting model to a certain level ac-ceptable for solving a forecasting problem in terms of accuracy and computational complexity.

Various drift formulas of the concept based on conditional and unconditional probabilities of obtaining the target variable depending on feature vector data in input dataset are studied. The notions of positive and negative drift of the concept are introduced depending on their belonging to the corresponding class used in forecasting.

A new approach uses a conceptual knowledge base of subject area, which allows a priori classifying the elements of feature vector in the form of generic taxonomy. A classified feature vector is a hierarchical structure that allows using bootstrap algorithm to form subsamples of fea-tures (folds) for preliminary training of weak models of the first and second levels. In this case, the folds can be ordered and then a resulting order can be used to identify and compensate for co-variant drift of working model concept.

As an example for experimental research, the subject area of air transportation services from the Kaggle international repository is taken. Software implementation was performed using Spider v.4 toolkit in Python v.4. The results of the experiments show the effectiveness of a new approach to correct concept drift..

***The aim of the work** is to obtain a new approach to identify and correct a covariant concept drift, which makes it possible to correct the drift in ensembles of machine learning models.*

***Keywords:** concept drift, ensembles of machine learning models, big data, prediction accuracy, knowledge base, ontological knowledge models.*

DOI: 10.21667/1995-4565-2023-85-44-53

References

1. **M. Asghari, S. Sierra-Sosa, M. Telahun, A. Kumar, A. S. Elmaghraby,** Aggregate density-based concept drift identification for dynamic sensor Data models, Neural *Comput. Appl.* 33 (8) (2021), pp. 3267-3279.

2. **Kashirin I. Yu.** Data Mining s ispol'zovaniyem iyerarkhicheskikh chisel v retrospektivnoy diagnostike. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2022, no. 79, pp. 118-126. (in Russian).

3. **Schröerab Ch., Kruseb F., Gómezb J. M.** A Systematic Literature Review on Applying CRISP-SM Process Model, *ProceSia Computer Science* 181 (2021), pp. 526-534.

4. **Breiman L.** Bagging predictors. *Machine Learning*, 24:123-140, 1996.

5. **Tsiakmaki M., Kostopoulos G., Kotsiantis S., Ragos O.** Transfer learning from deep neural networks for predicting student performance. *Applied Sciences*, vol. 10, no. 6 (2020) pp. 2145–2218.

6. **Bayram F., Ahmed B. S., Kassler A.** From concept drift to model degradation: An overview on performance-aware drift detectors. *Knowledge-Based Systems* 245 (2022) 108632.

7. **Ali T. M., Nawaz A., Rehman A. et al.** A sequential machine learning-cum-attentions mechanism for effective segmentation of brain tumor. *Frontiers in Oncology*, vol. 12, pp. 1-10, 2022.

8. **Tsai F. S., Cabrilo S., Chou H. H., Hu F. , Tang A. S..** Open innovation and SME performance: the roles of reverse knowledge sharing and stakeholder relationships. *Journal of Business Research*, vol. 148, pp. 433-443, 2022.

9. **Xia K., Kai-Zhan Lee, Bengio Y., Bareinboim E.** The causal-neural connection: Expressiveness, learnability, and inference. *Advances in Neural Information Processing Systems*, 34:10823-10836, 2021.

10. **Wang T. S., Parsia B., Hendler J.** A Survey of the Web Ontology Landscape. *The Semantic Web – ISWC 2006*. Lecture Notes in Computer Science, vol. 4273, 682 p.

11. **Kashirin I. Yu., Filatov I. Yu**. FormalizeS Sescription Of Intuitive Perception Of Spatial Situations. *2019 8th MeSiterranean Conference on Embedded Computing (MECO),* Budva, Montenegro, 2019, pp. 1-4.

12. **Kashirin I. Yu.** Neyronnyye seti, ispol'zuyushchiye modeli znaniy. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2021, no. 75, pp. 71-84. (in Russian).

13. **Satybaldina D. Zh., Ovechkin G. V., Kalymova K. A.** Sistema raspoznavaniya staticheskikh zhestov ruk s ispol'zovaniyem kamery glubiny. *Vestnik ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta.* Ryazan', 2020, no. 72, pp. 93-105. (in Russian).

14. **Phuong T. Nguyen, Davide Di Ruscio, Alfonso Pierantonio, Juri Di Rocco, and Ludovico Iovino.** Convolutional neural networks for enhanced classification mechanisms of metamodels. *Journal of Systems and Software* 172 (2021).