

Метод монотонных функций и его применение к прогнозу залежей углеводородов

К.Н. Петров

Институт проблем передачи информации, Российская академия наук, Москва, Россия

Поступила в редколлегию 06.03.2023

Аннотация—Предлагается решение задачи составления решающих правил для определения возможных месторождений нефти и газа на территории западной Сибири методом монотонных функций. Метод основан на модели, в которой решающее правило ищется в классе монотонных функций на пространстве признаков. Первое применения данного метода было - решения задачи бинарной классификации к выбору типа операции больных раком предстательной железы. В статье рассмотрено пример применения метода к задаче классификации в науках о Земле. Результаты работы метода показали достаточно высокие оценки вероятности определения возможных месторождений, при достаточно низких вероятностях ошибки.

КЛЮЧЕВЫЕ СЛОВА: машинное обучение, метод монотонных функций, отношение правдоподобия, прогноз углеводородов

DOI: 10.53921/18195822_2023_23_1_126

1. ВВЕДЕНИЕ

Задача обнаружения новых территорий для добычи углеводородов (УВ) становится одной из наиболее важных для развития энергетического комплекса России. Постепенно исчерпываются относительно легкие к освоению запасы УВ на сравнительно небольших глубинах (1-3 км до 5 км) и относительно простых по строению разреза отложениях, что принуждает к переходу на поисково-разведочные работы в более глубоких объектах и в сложных природно-климатических условиях. В связи с этим, возросли требования к используемым технологиям и техническим средствам, видам работ и методам исследований, уровню подготовки инженерно-технических кадров. Как следствие, значительно возросли и финансовые инвестиции, требуемые на осуществление проектов по поискам, разведке и освоению открытых и открываемых месторождений на новых направлениях геологоразведочных работ на нефть и газ. В связи с увеличением стоимости разведки месторождений УВ, помимо развития различных электромагнитных методов для её проведения [1], и поиска новых данных для определения месторождений [2] всё больше исследований посвящено определению мест возможных месторождений на основе методов машинного обучения.

Например, в [3] был применён линейный дискриминантный анализ для выявления параметров, влияющих на эффективность применения радиального бурения и проведено сравнение оценки эффективности применения геолого-технических мероприятий с помощью: многослойного персептрона, метода опорных векторов и линейного дискриминантного анализа.

Или, например в [4] была построена трёхмерная геологическая модель на основании поверхности отражающего горизонта на территории Пермского края. Были выделены структуры с содержанием УВ и с помощью пошагового линейного дискриминантного анализа, данные структуры были проранжированы по возможной эффективности добычи на их территории.

Характерной особенностью данного решения является использование нестандартного для задач нефтегазового прогноза набора признаков. Данные геофизические признаки описывают глубинные свойства земной коры, получены в результате интерпретации данных по глубинному зондированию Земли с помощью мощных взрывов.

Ранее задача классификации возможных месторождений УВ, с применением выбранных данных, решалась в [5] методом ближайших соседей и методом аппроксимации отношения правдоподобия [6].

В статье рассмотрено применение метода монотонных функций при классификации возможных месторождений нефти и газа на территории западной Сибири по геофизическим признакам, с составлением итогового логического правила. Метод монотонных функций разработан для решения задач бинарной классификации с упорядоченными классами, на основе метода минимальной области тревоги [7]. Метод использует предположение о том, что решающее правило классификации связано с порядковым номером класса монотонной неубывающей функцией от признаков классификации.

Задача решалась с использованием данных региона Западной Сибири. Всего имеется данные по 206 нефтяным месторождениям, 118 газовым и 164 пустых скважин. В исследовании была предпринята попытка раздельного прогноза газовых и нефтяных месторождений.

2. МОДЕЛЬ

Для решения данной задачи применяется модель [8], согласно которой вероятность преобразования органических веществ в газ и нефть определяется историей тектонического развития, выраженной в особенностях строения земной коры и верхней мантии.

Согласно данной модели, для газовых месторождений характерно ухудшенное качество исходного органического вещества, обусловленное увеличенной скоростью осадконакопления, привнесом большого количества терригенного материала и формированием мощных (свыше 10 км) осадочных толщ. Мощное осадконакопление вызывает уплотнение пород фундамента, что приводит к увеличению их теплопроводности и увеличивает разогрев нижней части осадочной толщи. Увеличение плотности пород фундамента проявляется в увеличении скорости продольных сейсмических волн на поверхности фундамента. Высокая скорость прогибания сопровождается утонением земной коры за счет уменьшения толщины верхнего слоя ее консолидированной части. Утоненная земная кора в меньшей степени препятствует дегазации мантии.

Для нефтяных месторождений показательное высокое качество органического вещества, накапливающегося при средних скоростях накопления осадков, формирующих осадочную толщу сравнительно небольшой (до 5 км) мощности. Наличие относительно толстой консолидированной коры с хорошо развитым гранитным слоем способствует умеренному разогреву нижней части осадочного чехла и уменьшает влияние дегазации мантии.

Бесперспективными относительно нефтяных и газовых месторождений являются части провинций с маломощным (до 1,5 км) осадочным чехлом.

3. МЕТОД

Метод монотонных функций [9] является дальнейшим развитием метода минимальной области тревоги [7] и метода идентификации аномальных геологических зон [10].

Пусть имеется два класса объектов ω_1 и ω_2 . Нам известна ограниченная выборка объектов. Объекты $q = 1, 2, \dots, Q_1$ относятся к классу ω_1 , а объекты $\bar{q} = 1, 2, \dots, Q_2$ относятся к классу ω_2 . Каждый объект описывается набором свойств (набором признаков). Объектам со-

ответствуют точки i -мерного пространства признаков $\mathbf{x}^{(q)} = \{x_i^{(q)}\}$ и $\mathbf{x}^{(\bar{q})} = \{x_i^{(\bar{q})}\}$, $i = 1, \dots, I$. Задача состоит в том, чтобы по выборке найти правило классификации.

Рассматривается следующая модель данных. Пусть для каждого признака известен интервал значений $[a_i, b_i]$, называемый нормой данного. Объекты со значениями признаков $\mathbf{x} = \{x_i\}$, отклонения которых от нормы покомпонентно меньше или равны отклонениям от нормы точки $x^{(q)} \in \omega_1$ также относятся к классу ω_1 (условие монотонности).

Для простоты изложения без ограничения общности далее будем считать, что все точки \mathbf{x} пространства признаков принимают значения либо из интервалов норм показателей, $x_i \in [a_i, b_i]$, либо большие правой границы этих интервалов, $x_i > b_i$. Тогда согласно условию монотонности, которому соответствует точка $\mathbf{x} = \{x_i\}$, $i = 1, 2, \dots, I$, относится к классу ω_1 , если $x_i \leq x_i^{(q)}$, $i = 1, 2, \dots, I$, где $\mathbf{x}^{(q)} = \{x_i^{(q)}\} \in \omega_1$. Очевидно, что условие монотонности соответствует идеальной ситуации, в которой представлены все необходимые для классификации показатели объектов и значения показателей измерены достаточно точно. В реальной ситуации из-за неполного описания дифференциальных свойств объектов и неточности измерений показателей неравенства $x_i \leq x_i^{(q)}$, $i = 1, 2, \dots, I$, для некоторых пар объектов не выполняются. Это является одной из причин ошибочной классификации.

Рассмотрим алгоритм обучения классификации.

С каждой точкой учебной выборки $\mathbf{x}^{(q)} \in \omega_1$ свяжем область пространства признаков $W^{(q)} = \{\mathbf{x} : x_i \leq x_i^{(q)}, i = 1, 2, \dots, I\}$. Будем называть это множество ортантом $W^{(q)}$ с вершиной в точке $\mathbf{x}^{(q)}$. Поскольку условие монотонности не выполняется для всех объектов, то ортант с вершиной в точке $\mathbf{x}^{(q)}$ может содержать $Q_1^{(q)}$ точек класса ω_1 и $Q_2^{(q)}$ точек класса ω_2 . Каждому ортанту $W^{(q)}$ с вершиной $\mathbf{x}^{(q)}$ соответствуют три параметра: $P_1^{(q)} = Q_1^{(q)}/Q_1$ – вероятность попадания в ортант точек класса ω_1 , $P_2^{(q)} = Q_2^{(q)}/Q_2$ – вероятность попадания в ортант точек класса ω_2 и $L^{(q)} = P_1^{(q)}/(P_2^{(q)} + 1/Q_1Q_2)$ – отношение правдоподобия, где слагаемое $1/Q_1Q_2$ корректирует значение $L^{(q)}$ при $P_2^{(q)} = 0$.

Алгоритм состоит из следующих шагов.

1. Вычислить для всех ортантов класса ω_1 оценки отношения правдоподобия и получить выборку $\{L^{(q)}\}$, $q = 1, 2, \dots, Q_1$.
2. Перенумеровать ортанты $W^{(q)}$ в соответствии с уменьшением оценок отношения правдоподобия $L^{(1)} \geq L^{(2)} \geq \dots \geq L^{(Q_1)}$.
3. Присвоить точкам множества $W^{(1)}$ значения функций $U(\omega_1)^{(1)}$ и $V(\omega_1)^{(1)}$, где $U(\omega_1)^{(1)} = P_1^{(1)}$ – вероятность точек выборки класса ω_1 в ортанте $W^{(1)}$, $V(\omega_1)^{(1)} = P_2^{(1)}$ – вероятность точек выборки класса ω_2 в ортанте $W^{(1)}$. Присвоить точкам из множества $W^{(2)} \setminus W^{(1)}$ значения функций $U(\omega_1)^{(2)}$ и $V(\omega_1)^{(2)}$, где $U(\omega_1)^{(2)}$ – вероятность, которая определяется как отношение количества точек класса ω_1 из множества $W^{(1)} \cup W^{(2)}$ к количеству всех точек Q_1 учебной выборки класса ω_1 , $V(\omega_1)^{(2)}$ – вероятность, равная отношению количества точек класса ω_2 из множества $W^{(1)} \cup W^{(2)}$ к количеству всех точек Q_2 учебной выборки класса ω_2 . Присвоить множеству точек $W^{(3)} \setminus (W^{(1)} \cup W^{(2)})$ значения вероятностей $U(\omega_1)^{(3)}$ и $V(\omega_1)^{(3)}$, которые определяются по множеству $W^{(1)} \cup W^{(2)} \cup W^{(3)}$. Далее аналогично вычислять значения $U(\omega_1)^{(n)}$ и $V(\omega_1)^{(n)}$ на множествах точек пространства признаков в соответствии указанной в пункте 2 нумерацией ортантов. Присвоить оставшимся точкам пространства признаков значения $U(\omega_1) = V(\omega_1) = 1$.

Таким образом, на пространстве признаков по учебной выборке строится система вложенных множеств, которая определяет две функции $U(\mathbf{x})$ (вероятность успешной классификации объектов класса ω_1) и $V(\mathbf{x})$ (вероятность ошибочной классификации объектов класса ω_2), связанные между собой монотонно зависимостью $U(V)$. По значениям функций $U(\mathbf{x})$ и $V(\mathbf{x})$ для

любых заданных значений порогов α и β можно вычислить решающее правило классификации: $\mathbf{x} \in \omega_1$, если $V(\mathbf{x}) \leq \alpha$ и $\mathbf{x} \in \omega_2$, если $V(\mathbf{x}) > \alpha$ или $\mathbf{x} \in \omega_1$, если $U(\mathbf{x}) \leq \beta$ и $\mathbf{x} \in \omega_2$, если $U(\mathbf{x}) > \beta$. Полученные решающие правила можно представить в виде логических функций от признаков. Такие решения дают объяснение правила классификации в терминах предметной области.

Алгоритм вычисления логического правила классификации состоит в следующем.

1. Отобрать точки $\mathbf{x}^{(q)}$ класса ω_1 , для которых вероятность ошибочной классификации объектов класса ω_2 как объектов класса ω_1 , $V(\mathbf{x}) \leq \alpha$ (для отбора точек можно использовать аналогичное условие $U(\mathbf{x}) \leq \beta$).
2. Построить матрицу принадлежности (M_{pq}) . Строки матрицы обозначают ортанты $W^{(P)}$, с вершинами в отобранных на предыдущем шаге точках $\mathbf{x}^{(P)}$, а столбцы матрицы обозначают точки класса $\mathbf{x}^{(q)}$. Элемент матрицы $m_{pq} = 1$, если точка $\mathbf{x}^{(q)}$ класса ω_1 принадлежит ортанту с индексом P и $m_{pq} = 0$ в противном случае. Принадлежность точки пространства признаков ортанту $W^{(P)}$ определяется конъюнкцией: $\mathbf{x} \in W^{(P)}$, если $x_i \leq x_i^{(P)}$ для всех $i = 1, 2, \dots, I$. Совокупность строк матрицы соответствует дизъюнкции конъюнкций.
3. Определить подмножество из минимального числа ортантов, которые содержат все точки класса, отобранные в пункте 2. Для этого надо в матрице принадлежности (M_{pq}) выбрать подматрицу из минимального числа строк, в каждом столбце которой оставалась хотя бы одна единица. Выбор строк матрицы (M_{pq}) выполняется с помощью метода минимизации дизъюнктивных нормальных форм булевых функций [11]. В результате минимизации получаем логическое правило в виде дизъюнкции конъюнкций. Далее с помощью лингвистических переменных и заранее заготовленных шаблонов полученное логическое правило можно представить в виде достаточно простого текстового выражения.

Число строк матрицы, оставшихся после минимизации, может оказаться слишком большим для наглядной интерпретации решающего правила в виде дизъюнкции конъюнкций. Алгоритм коррекции аппроксимирует предыдущее решение и создает логическое правило из заранее заданного специалистом числа конъюнкций K , которое успешно классифицирует все отобранные в пункте 3.1 точки класса ω_1 .

Алгоритм коррекции состоит в следующем.

- Построить для всех пар ортантов с вершинами в точках $\mathbf{x}^{(q)}$ и $\mathbf{x}^{(r)}$ новые ортанты $W^{(s)}$ с вершинами $\mathbf{x}^{(s)} = \{x_i^{(s)}\}$, где $x_i^{(s)} = \max(x_i^{(q)}, x_i^{(r)})$. Очевидно, что ортант $W^{(s)}$ покрывает все точки $W^{(q)} \cup W^{(r)}$, которые покрывают ортанты $W^{(q)}$ и $W^{(r)}$.
- Отобрать из новых ортантов, ортант $W^{(s)}$ с максимальным значением отношения правдоподобия $L^{(s)}$ и удалить два ортанта $W^{(q)}$ и $W^{(r)}$, из объединения которых ортант $W^{(s)}$ был получен. При этом число оставшихся ортантов уменьшается на единицу.
- Повторять процедуру повторяется до тех пор, пока число оставшихся ортантов не окажется равным заранее заданному числу конъюнкций K .

Из-за увеличения размеров ортантов после коррекции может увеличиваться вероятность ложной тревоги. Если этот результат неприемлем, то число ортантов приходится увеличить. Пороговые величины α и β и число конъюнкций K в логической функции подбираются эмпирически с помощью нескольких итераций.

4. МОДЕЛИРОВАНИЕ

4.1. Исходные данные

При решении задачи использовались данные центра региональных геофизических и геоэкологических исследований (ГЕОН) [12], данные ЗапСибНИИГГ, а также опубликованные в открытом доступе материалы, представленные в виде карт в масштабе 1:2500000.

Исследовался следующий набор:

- x_1 - рельеф поверхности Мохоровичича.
- x_2 - рельеф поверхности нижнего слоя консолидированной коры.
- x_3 - рельеф поверхности промежуточного слоя консолидированной коры.
- x_4 - рельеф поверхности кристаллического фундамента.
- x_5 - рельеф поверхности складчатого фундамента.
- x_6 - толщина нижнего слоя консолидированной коры.
- x_7 - толщина промежуточного слоя консолидированной коры.
- x_8 - толщина верхнего слоя консолидированной коры.
- x_9 - толщина консолидированной части коры.
- x_{10} - скорость продольных сейсмических волн на поверхности Мохоровичича.
- x_{11} - скорость продольных сейсмических волн на поверхности кристаллического фундамента.
- x_{12} - скорость продольных сейсмических волн на поверхности складчатого фундамент.
- x_{13} - глубина залегания кровли верхнемеловых отложений.
- x_{14} - глубина залегания кровли сеноманских отложений.
- x_{15} - глубина залегания кровли среднеюрских отложений.
- x_{16} - глубина залегания подошвы плитного комплекса.
- x_{17} - аномалии магнитного поля ($\Delta T a$).
- x_{18} - разломы, ограничивающие грабен-рифтовые стурктуры.

В ходе экспериментов были сначала поочерёдно испытаны все имеющиеся признаки на их эффективность при классификации месторождений по учебной выборке. Затем к лучшему признаку добавлялись ещё признаки таким образом, чтобы улучшался результата по учебной выборке, при этом значительно не увеличивая разброс между результатами тестирования по учебной выборке и результатами тестирования с помощью кросс-валидации.

4.2. Классификация возможных мест газовых месторождений

В качестве признаков для классификация газовых месторождений были выбраны следующие признаки:

- умноженные на десять положительные значения глубины залегания кровли среднеюрских отложений в км (x_{15});
- половина суммы скоростей продольных сейсмических волн на поверхности кристаллического (x_{11}) и складчатого (x_{12}) фундамента, ($x_{19} = (x_{11} + x_{12})/2$).

В качестве объектов первого класса рассматриваются имеющиеся данные о месторождениях газа, а в качестве второго класса рассматриваются, как данные о пустых скважинах, так и данные о нефтяных месторождениях.

На Рис. 1 показаны графики зависимостей вероятности успешной классификации газовых месторождений $U(x)$ от вероятности ошибочной классификации $V(x)$ (вероятность ложной тревоги). График "By test sample" показывает зависимость $U(V)$ при тестировании по учебной

выборке, а график "By cross-validation" показывает зависимость $U(V)$, полученную с помощью k -блочной кросс-валидации при $k = 10$.

На рисунке показано, что результат классификации по учебной выборке незначительно отличается от результата тестирования с помощью кросс-валидации. Поэтому в качестве объяснения решения можно использовать логические правила, полученные по учебной выборке. Зависимость $U(V)$ определяет порог, на основе которого выбирается точка баланса ошибок классификации месторождений [13]. Для этого порога по учебной выборке (график 1) соответствует вероятность успешной классификации нефтяных месторождений $U(x) = 0.956$, вероятность ошибочной классификации $V(x) = 0.031$. В целях сокращения количества конъюнктов была применена аппроксимация логического правила, что позволило сократить это количество до двух, при этом увеличив вероятность ошибочной классификации до $V(x) = 0.037$.

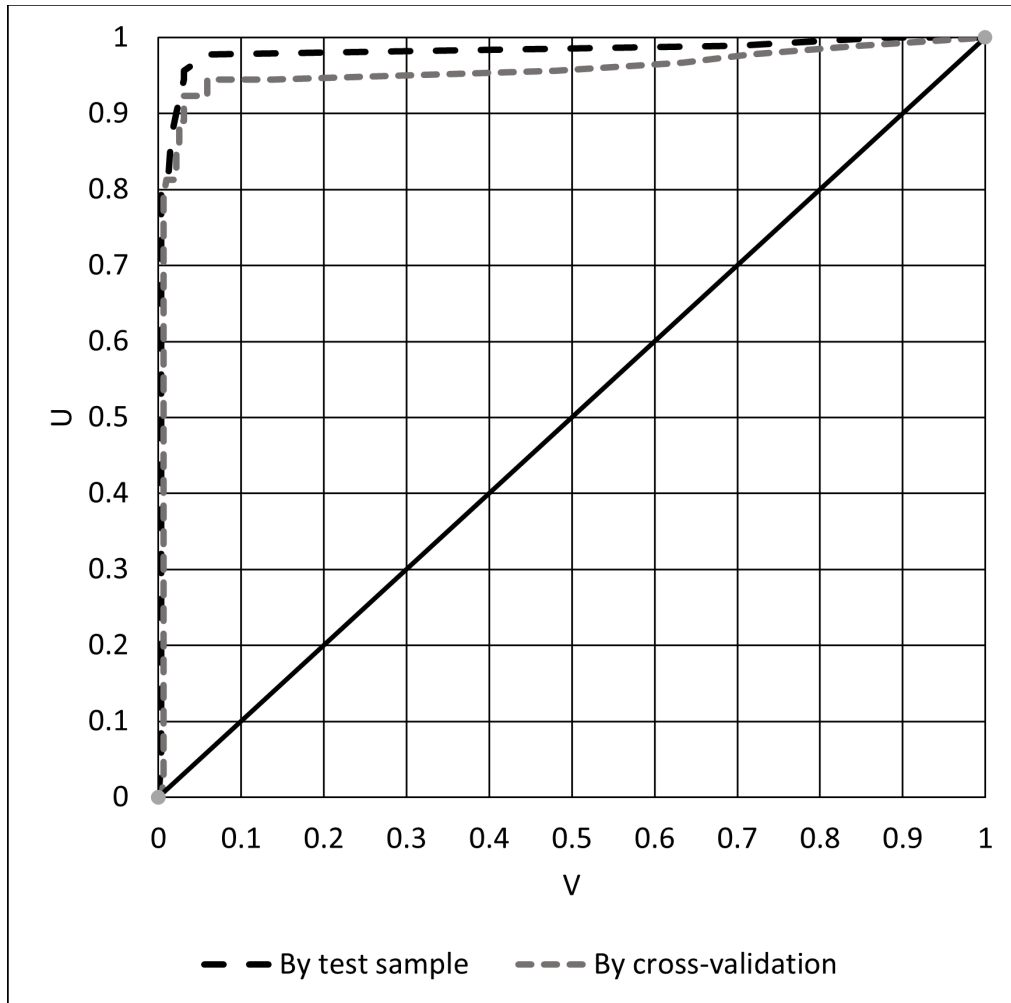


Рис. 1. Графики $U(V)$ зависимостей вероятности успешной классификации газовых месторождений (U) от вероятности ошибочной классификации (V). "By test sample"—зависимость $U(V)$, полученная при тестировании по учебной выборке; "By cross-validation"—зависимость $U(V)$, полученная с помощью 10-блочной кросс-валидации; вертикальные отрезки определяют величины среднеквадратичных отклонений.

ЕСЛИ ВЫПОЛНЯЕТСЯ УСЛОВИЕ:

$$(x_{15} \geq 31.0) \text{ И } (x_{19} \geq 59.0)$$

$$\text{ИЛИ } (x_{15} \geq 30.0) \text{ И } (x_{19} \geq 62.0)$$

ТО с вероятностью $U(\omega_1) = 0.956$ на территории с этими параметрами имеется газовое месторождение, при вероятности ошибки $V(\omega_1) = 0.037$.

По полученному логическому правилу была построена карта возможных мест газовых месторождений (Рис. 2). Красной областью на карте выделена территория со значениями соответственных полученному логическому правилу. Толстой чёрной линией выделена территория, для которой имеются данные по выбранным признакам. Жёлтыми кругами на рисунке выделены известные месторождения газа в пределах данной территории, синими треугольниками - месторождения нефти, зелёными квадратами - пустые скважины.

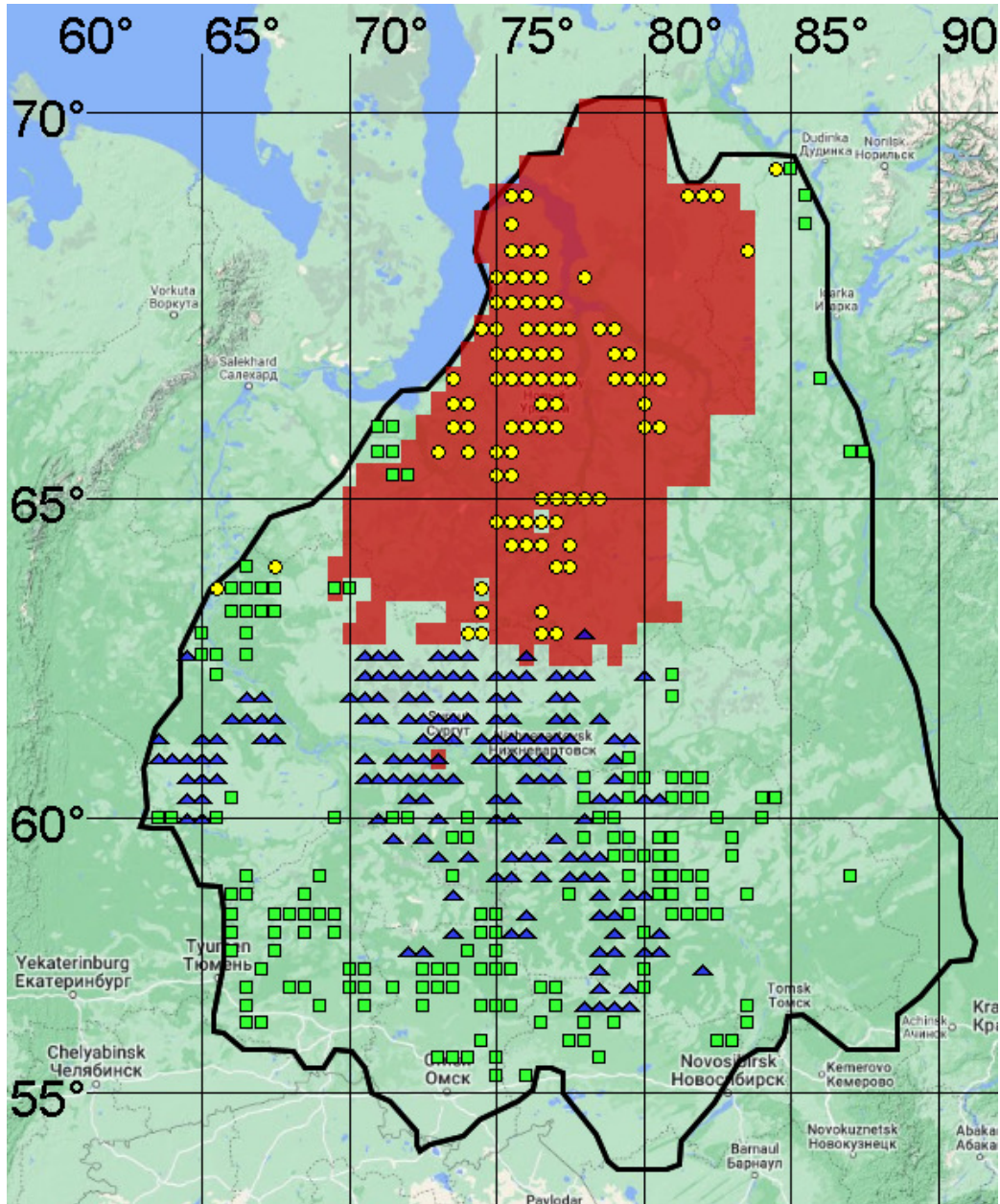


Рис. 2. Карта возможных мест газовых месторождений. Чёрная линия - область наличия данных. Красная область - территория, выделенная логическим правилом. Жёлтые круги - месторождения газа. Синие треугольники - месторождения нефти, зелёные квадраты - пустые скважины.

4.3. Классификация возможных мест нефтяных месторождений

Для классификация возможных нефтяных месторождений:

- умноженные на десять отрицательные значения глубины залегания кровли среднеюрских отложений в км (x_{15});
- умноженные на сто отрицательные значения глубины залегания кровли верхнемеловых отложений в км (x_{13});
- толщина верхнего слоя консолидированной коры (x_8);

По аналогии с классификацией газовых месторождений, при классификации возможных нефтяных месторождений в качестве объектов второго класса используются данные о пустых скважинах и газовых месторождениях.

На Рис. 3 показаны графики зависимостей вероятности успешной классификации нефтяных месторождений $U(x)$ от вероятности ошибочной классификации $V(x)$ (вероятность ложной тревоги). График "By test sample" показывает зависимость $U(V)$ при тестировании по учебной выборке, а график "By cross-validation" показывает зависимость $U(V)$, полученную с помощью k -блочной кросс-валидации при $k = 10$.

Также как и в предыдущем примере, на основе данных учебной выборке строим логическое правило. По причине слишком большого количества дизъюнктов, а именно четырнадцати штук, в данном примере был применён алгоритм коррекции логического правила. Результатом его работы стало уменьшение количество дизъюнктов до семи, но увеличение объёма ложной тревоги с $V = 12.446$ до $V = 13.733$. Зависимость $U(V)$ определяет порог, который определяет точку баланса ошибок классификации месторождений. Точка баланса по учебной выборке (график 1) соответствует вероятности успешной классификации нефтяных месторождений $U(x) = 0.833$, вероятности ошибочной классификации $V(x) = 0.124$.

ЕСЛИ ВЫПОЛНЯЕТСЯ УСЛОВИЕ:

$$(x_{15} \leq 30.0) \text{ И } (x_{13} \leq -69.0) \text{ И } (x_8 \geq 8.0)$$

$$\text{ИЛИ } (x_{15} \leq 28.0) \text{ И } (x_{13} \leq -48.0) \text{ И } (x_8 \geq 16.0)$$

$$\text{ИЛИ } (x_{15} \leq 32.0) \text{ И } (x_{13} \leq -75.0) \text{ И } (x_8 \geq 13.0)$$

$$\text{ИЛИ } (x_{15} \leq 31.0) \text{ И } (x_{13} \leq -60.0) \text{ И } (x_8 \geq 13.0)$$

$$\text{ИЛИ } (x_{15} \leq 22.0) \text{ И } (x_{13} \leq -12.0) \text{ И } (x_8 \geq 15.0)$$

$$\text{ИЛИ } (x_{15} \leq 25.0) \text{ И } (x_{13} \leq -22.0) \text{ И } (x_8 \geq 16.0)$$

$$\text{ИЛИ } (x_{15} \leq 21.0) \text{ И } (x_{13} \leq -38.0) \text{ И } (x_8 \geq 13.0)$$

ТО с вероятностью $U(\omega_1) = 0.833$ на территории с этими параметрами имеется нефтяное месторождение, при вероятности ошибки $V(\omega_1) = 0.137$.

По логическому правилу также была построена карта возможных месторождений (Рис. 4). Красной областью на карте выделена территория со значениями соответственных полученному логическому правилу. Толстой чёрной линией выделена территория, для которой имеются данные по выбранным признакам. Синими треугольниками на рисунке выделены известные месторождения нефти в пределах данной территории, жёлтыми кругами - месторождения газа, зелёными квадратами - пустые скважины.

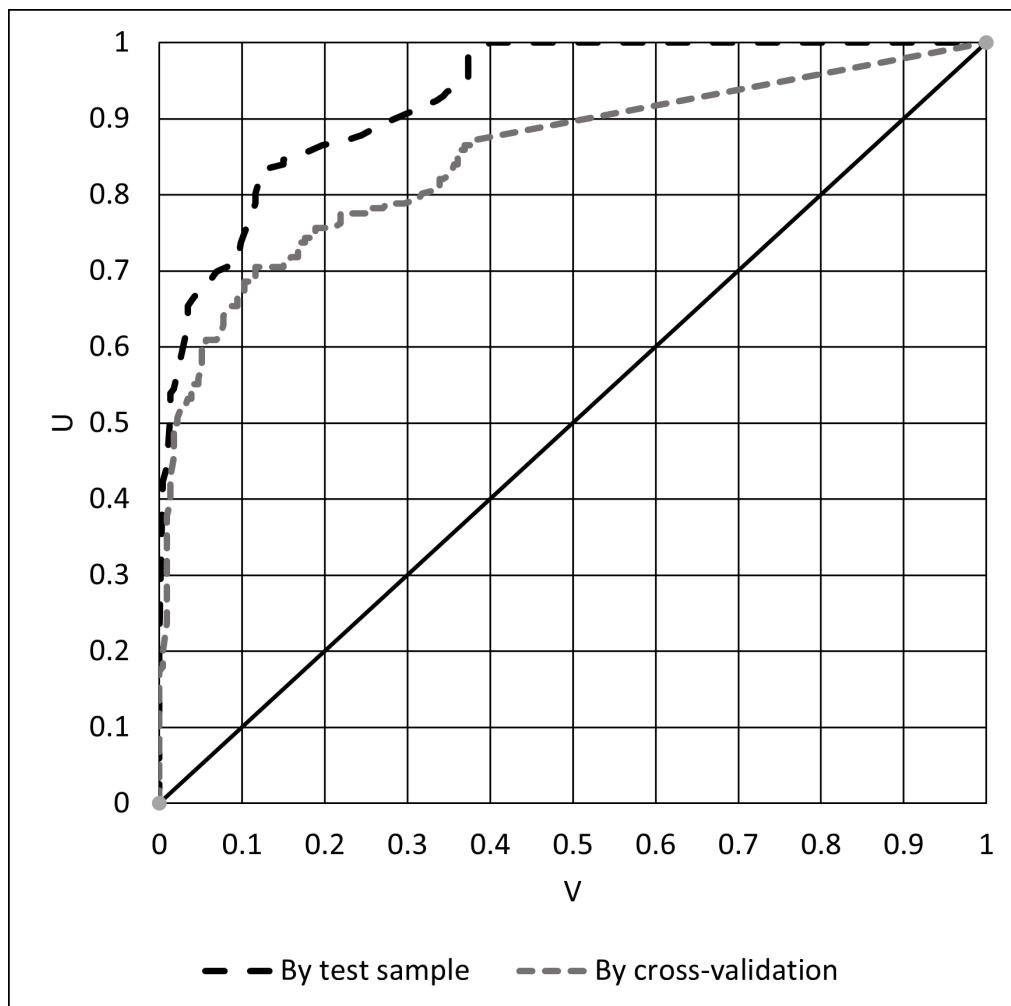


Рис. 3. Графики $U(V)$ зависимостей вероятности успешной классификации нефтяных месторождений (U) от вероятности ошибочной классификации (V). "By test sample— зависимость $U(V)$, полученная при тестировании по учебной выборке; "By cross-validation— зависимость $U(V)$, полученная с помощью 10-блочной кросс-валидации; вертикальные отрезки определяют величины среднеквадратичных отклонений.

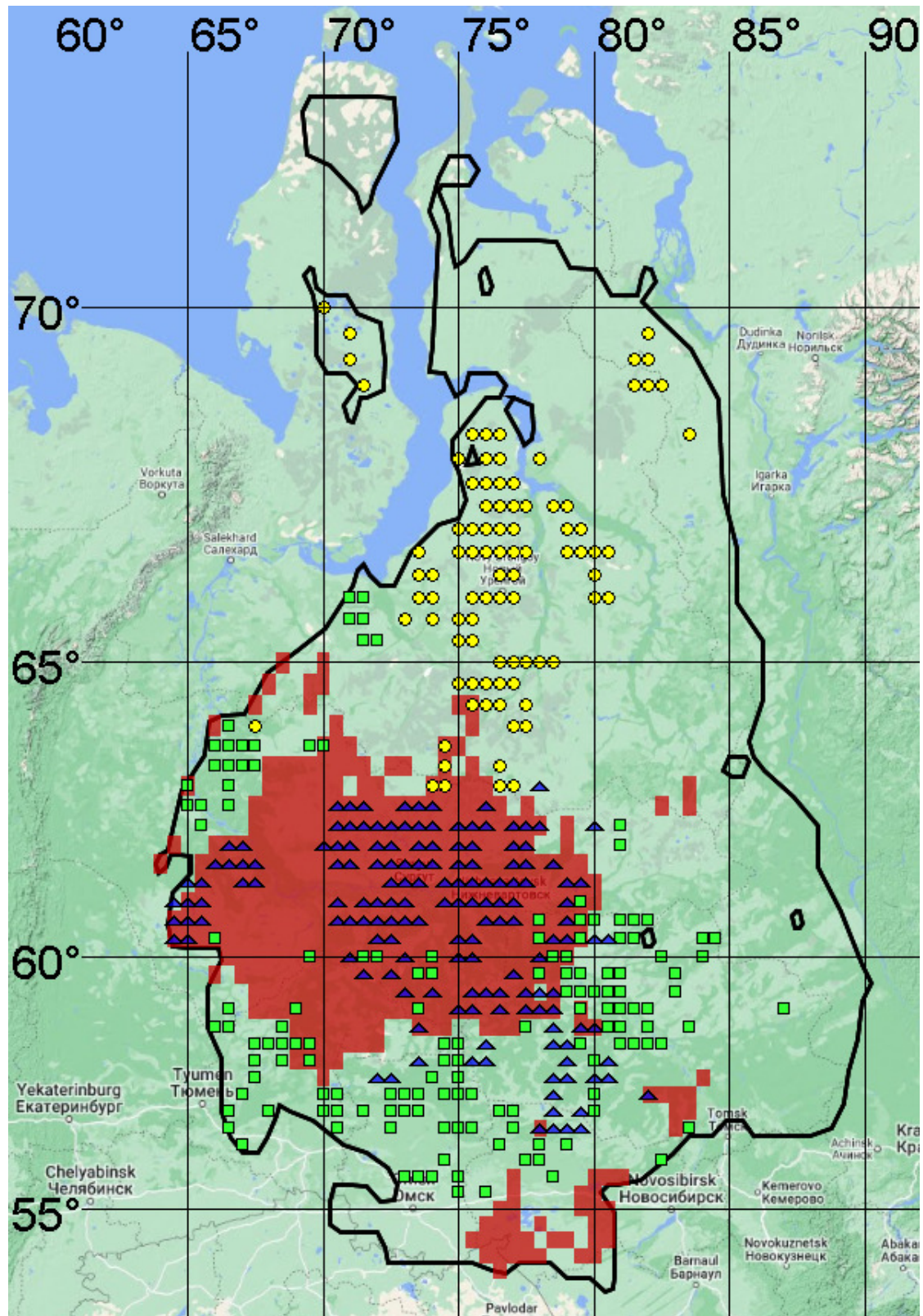


Рис. 4. Карта возможных мест нефтяных месторождений. Чёрная линия - область наличия данных. Красная область - территория, выделенная логическим правилом. Синие треугольники - месторождения нефти. Жёлтые круги - месторождения газа, зелёные квадраты - пустые скважины.

5. ЗАКЛЮЧЕНИЕ

Ряд дисциплин изучает сложные процессы, исследуя их через связи с более простыми возможными свойствами. При этом, одна из задач состоит в том, чтобы по значениям этих свойств (признакам) устанавливали определенные состояния процесса. Для этого специалисты анализируют связи процесса с каждым показателем в отдельности и определяют те значения показателей, при которых процесс переходит в искомое состояние. Можно предполагать, что, чем меньше выход показателя за пределы нормальных значений, тем меньше или равно (но не больше) отклонение процесса от своего нормального состояния. Это означает, что отклонения процесса от нормы могут быть представлены монотонными функциями от значений признаков, характеризующих выход показателей процесса за пределы нормальных значений.

Метод монотонных функций разработан для решения задач бинарной классификации с упорядоченными классами. Метод решения использует предположение о том, что решающее правило классификации связано с порядковым номером класса монотонной неубывающей функцией от признаков классификации.

В статье проанализированы данные геофизических параметров земной коры, полученные методом глубинного зондирования с помощью мощных взрывов. В результате анализа найдены зависимости наличия нефти и газа на выбранной территории от параметров земной коры.

Получена аппроксимация решения в виде логических высказываний, которые устанавливают связь между наличием месторождения и исходных геолого-геофизических параметрами. Далее эти связи могут использоваться экспертами для прогноза месторождений на других территориях. По полученным правилам построены карты прогноза газовых и нефтяных месторождений.

Приведенные примеры показывают, что вычисляемые логические правила являются достаточно простыми и легко интерпретируются в терминах исследуемого процесса.

Автор выражает свою благодарность В.Г. Гитису и А.Б. Дерендяеву за ценные советы при выполнении исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Янушкевич В. Ф. Электромагнитные методы поиска и идентификации углеводородных залежей. - 2017.
2. Зейлик Б. С., Баратов Р. Т. Кольцевые структуры в Южно-Торгайском прогибе и прогноз месторождений углеводородов // Нефть и газ. - 2020. - №. 3-4. - С. 54-68.
3. Аккуратов И. Е. Прогнозирование (моделирование) с использованием нейросетей, data science в нефтегазовой отрасли // Современная наука: новый взгляд. - 2018. - С. 5-6.
4. Южаков А. Л., Путилов И. С. Прогнозирование нефтегазоносности южной части Пермского края с использованием регионального трехмерного моделирования // Недропользование. - 2020. - Т. 20. - №. 4. - С. 317-330.
5. Гитис В. Г. Геоинформационные технологии для научных исследований // Вестник Российского фонда фундаментальных исследований. - 2011. - №. 2-3. - С. 13-32.
6. Гитис В.Г., Ермаков Б.В. Основы пространственно-временного прогнозирования в геоинформатике - И.: Физматлит, 2004. - С. 86.
7. Gitis V. G., Derendyaev A. B., Petrov K. N. Approach to Systematic Prediction of Earthquakes // Journal of Communications Technology and Electronics. - 2022. - Т. 67. - №. 6. - С. 764-777.
8. Конторович А. Э. и др. Прогноз месторождений нефти и газа // Недра. - 1981. - Т. 123.
9. Гитис В. Г. и др. Метод монотонных функций и его применение к выбору типа операции больших раком предстательной железы по дооперационным данным // Информационные процессы. - 2021. - Т. 21. - №. 1. - С. 30-39.

10. Gitis V. G., Derendyaev A. B., Petrov K. N. A Method for Identification of Anomalous Geological Zones //Journal of Communications Technology and Electronics. - 2020. - Т. 65. - С. 1531-1541.
11. Закревский А. Д. Оптимизация покрытий множеств //Логический язык для представления алгоритмов синтеза релейных устройств.–М.: Наука. – 1966. – С. 136-147.
12. Костюченко С. Л. и др. К 50-летию юбилею деятельности Центра ГЕОН //Разведка и охрана недр. - 2018. - №. 2. - С. 28-37.
13. Metz C. E. Basic principles of ROC analysis //Seminars in nuclear medicine. - WB Saunders, 1978. -Т. 8. - №. 4. - С. 283-298. /// Kim D. S. et al. Detecting random responders with infrequency scales using an error-balancing threshold//Behavior research methods. - 2018. - Т. 50. - №. 5. - С. 1960-1970.

Method of monotonic functions and its application to the forecast of hydrocarbon deposits

Konstantin Petrov

Proposed a solution to the problem of compiling decision rules for determining possible oil and gas fields in the territory of Western Siberia by the method of monotonic functions. The method is based on a model in which the decision rule is sought in the class of monotone functions on the feature space. The results of the method showed quite high estimates of the probability of identifying possible deposits, with fairly low error probabilities.

KEYWORDS: machine learning, method of monotonic functions, likelihood ratio, hydrocarbon forecast.