

Применение алгоритмов обучения в метеорологии для предсказания конвективной активности

G. SZALAY*, L. MOLNÁR**, O. GULYÁS**

Введение

Одной из целей нашей работы являлось исследование возможностей применения алгоритмов обучения или распознавания, посредством осуществления конкретных практических применений. Другой целью — в связи с предыдущем — являлось сравнение алгоритмов между собой с точки зрения надежности, скорости сходимости и других факторов, на основании достигнутых при решении практических задач результатов.

В настоящей работе показывается применение нескольких алгоритмов обучения, действующих по методу обучения с учителем, для задачи, связанной с предсказыванием погоды.

Метеорологическая задача относится к предсказыванию конвективной активности, являющейся мерой атмосферной грозовой деятельности.

В начале данной работы ради облегчения понимания дается краткий обзор использованных алгоритмов обучения и реализующих их программ для ЭВМ.

Эта работа является частью деятельности имеющей место в НИИ Связи (ТКИ), по областям диагностики посредством ЭВМ и распознавания образов.

Решение излагаемой задачи произведено в рамках сотрудничества между Центральным Метеорологическим Институтом и НИИ Связи.

Создание метеорологической модели и подготовка данных были произведены в Центральном Метеорологическом Институте, а теоретические исследования, связанные с алгоритмами, изготовление машинных программ и опыты посредством ЭВМ произвелись в НИИ Связи. Оценка результатов произведена сотрудниками обоих Институтов сообща.

Полученные до сих пор результаты показали, что метеорологическая модель и излагаемые алгоритмы пригодны для решения поставленной задачи.

Однако нашей конечной целью является внедрение процедур для ЭВМ, использующих алгоритмы обучения в оперативную метеослужбу, при дальнейшем их развитии и увеличении их эффективности.

1. О выборе алгоритмов

В распознавании образов известно множество различных вариантов алгоритмов обучения. Подробный обзор их находится в [1]—[4].

Для выбранной нами задачи наиболее подходящими являются алгоритмы действующие по методу обучения с учителем.

Первую группу алгоритмов, использованных при решении задачи, составляют алгоритмы, действующие методом потенциальных функций. Метод потенциальных функций является одним из наиболее основательно разработанных видов алгоритмов обучения.

Эвристическое введение его содержится в [5] и [14], а теоретические основы и наиболее известные разновидности изложены в [6] и [7].

В НИИ Связи, кроме исследования и реализации при помощи ЭВМ, занимались и обобщением этих алгоритмов и другими теоретическими исследованиями [8]—[11].

Выбор метода алгоритмов потенциальных функций сверх вышеуказанных мотивировался и тем, что эти алгоритмы уже были успешно применены в других областях метеорологии [19].

Вторым использованным алгоритмом являлся метод полиномиальной дискриминации [ПДМ]. Теоретические основы ПДМ содержатся в [12], а модифицированные и усовершенствованные варианты в [14].

Его связь с методом потенциальных функций показывается в [14] и [15]. Основным достоинством ПДМ является то, что он очень надежно классифицирует и в случае наличия относительно малой обучающей последовательности. В работе [13] показывается успешное применение ПДМ в кардиологии. Третий использованный алгоритм, реализующий простейший из видов многослойных сетей пороговых элементов, так называемый алгоритм "committee machine" (СМ). СМ способен образовать кусочно-линейную поверхность разделения.

Теоретические основы линейного и кусочно-линейного разделения, а также разные варианты этого алгоритма содержатся в [3] и [16]. Об успешном применении для предсказания осадков алгоритма кусочно-линейного разделения сообщается в [17].

В последующем рассмотрим общие черты алгоритмов обучения, действующих по методу обучения с учителем, далее коротко излагаются теоретические основы вышеуказанных трех алгоритмов.

2. Алгоритмы обучения действующие по методу обучения с учителем

Излагаемые нами алгоритмы принадлежат к группе алгоритмов разделения с помощью разделяющих функций.

Наши алгоритмы обучения служат для разделения подмножеств n -мерного Эвклидова пространства R^n . Без ограничения общности можем предполагать число разделяемых подмножеств равным двум. Соответственно пусть будут заданы множества A и B в основном пространстве X , для которых

$$\begin{aligned} A \cup B &\subset X, \\ A \cap B &= \emptyset. \end{aligned} \tag{2.1}$$

Разделяющей функцией (функцией разделения) называются такие функции $f(x)$, что

$$\text{sign } f(x) = \begin{cases} +1 & \text{при } x \in A, \\ -1 & \text{при } x \in B. \end{cases} \quad (2.2)$$

Материалом обучения или *обучающей последовательностью* называется множество точек $\{x^i \in X\}_1^\infty$, если (1) выбор каждой точки x^i производится независимо (2) по общему распределению вероятностей $p(x)$.

У алгоритмов обучения, действующих по методу обучения с учителем, к каждой точке обучающей последовательности принадлежит указание о том, к какому множеству следует отнести данную точку, и это называется *учением*. Обучение же представляет собой последовательность вероятностных переменных $\{\hat{x}^i\}_1^\infty$, где

$$\begin{aligned} \hat{x}^i &= f(x^i), \text{ или} \\ \hat{x}^i &= \text{sign } f(x^i). \end{aligned} \quad (2.3)$$

Пусть $\hat{\pi}_n$ и π_n являются совокупностями n первых элементов обучения и обучающей последовательности соответственно.

Задачей алгоритмов обучения является выработка такой функции $f_n(x)$, которая на основании обучающей последовательности и обучения «в некотором смысле» образует $f(x)$, т. е.

$$f_n(x) = f_n(x; \pi_n, \hat{\pi}_n) \cong f(x), \quad (2.4)$$

$$\text{где } \pi_n = \{x^1, x^2, \dots, x^n\},$$

$$\hat{\pi}_n = \{\hat{x}^1, \hat{x}^2, \dots, \hat{x}^n\}.$$

2.а. Метод потенциальных функций

При алгоритмах потенциальных функций предполагаем, что разделяющая функция допускает запись в виде:

$$f(x) = \sum_{i=1}^N c_i \varphi_i(x) \quad (2.а.1)$$

где $\{\varphi_i(x)\}_{i=1}^N$ является линейной системой независимых функций.

n -ое приближение разделяющей функции $f_n(x)$ образуется следующим рекурсивным путем:

$$f_n(x) = f_{n-1}(x) + r_n(x^n) \cdot K(x, x^n) \quad (2.а.2)$$

где определяемая в виде

$$K(x, y) = \sum_{i=1}^N \varphi_i(x) \cdot \varphi_i(y) \quad (2.а.3)$$

функция называется потенциальной функцией, а корректирующая переменная $r_k(x^k)$ имеет различный, описанный в [6] и [7] вид при каждом конкретном алгоритме.

В дальнейшем при

$$r_k^\theta(x^k) = \theta [f(x^k) - f_{k-1}(x^k)]$$

$$0 \leq \theta \leq 2 / \sup_x K(x, x) \quad (2. a. 4)$$

алгоритм называем тета-алгоритмом, а при

$$r_k^\gamma(x^k) = \gamma_k \cdot \text{sign} [f(x^k) - f_{k-1}(x^k)]$$

$$\gamma_k > 0, \quad \sum_1^\infty \gamma_k = \infty, \quad \sum_1^\infty \gamma_k^2 < \infty \quad (2. a. 5)$$

— гамма-алгоритмом.

Теоремы сходимости, относящиеся к алгоритмам содержатся в [7] и [9].

В сделанной нами программе, реализующей алгоритмы обучения по методу потенциальных функций для ЭВМ, выбрали разделяющую функцию в виде полинома.

Количество переменных и число степеней полинома ограничивается лишь емкостью памяти ЭВМ. Коэффициенты полинома вычисляются рекурсивным путем, поэтому нет необходимости сохранять в памяти точки обучения. Вследствие этого потребность в памяти в течении обучения остается постоянной и не зависит от длины обучающей последовательности.

Программа — подобно всем описываемым в настоящем докладе программам — написана на языке Алгол — 60 в машинной репрезентации GIER—ALGOL 4.

2.6. Метод полиномиальной дискриминации (ПДМ)

Сводка теоретических основ алгоритма содержится в [12]. Алгоритм ПДМ способен дать не только приближение разделяющей функции, но кроме того, дает и приближительную оценку распределения вероятностей точек обучения, принадлежащих к каждому отдельному классу образов.

Оценка вышеуказанных функций плотностей распределений вероятностей производится таким образом, что для каждой точки обучающей последовательности назначается так называемая функция интерполяции и вычисляется их среднее значение.

В случае одномерного основного пространства отношения иллюстрируются на рисунке I.

С целью облегчения вычислительной работы в качестве интерполяционных функций были выбраны экспоненциальные функции.

Таким образом принадлежащая к классу A функция плотности принимает вид:

$$f_A(x) = \frac{1}{(\sigma\sqrt{2\pi})^p} \frac{1}{m} \sum_{i=1}^m \exp \left[- \sum_{k=1}^p (x_{aik} - x_k)^2 / 2\sigma^2 \right], \quad (2.6.1)$$

где:

m — число точек обучающей последовательности, принадлежащих к классу A ,
 p — число мер основного пространства,

$x_{ai} = (x_{ai1}, x_{ai2}, \dots, x_{aip})$ — i -тая точка обучающей последовательности.

При формировании правила решения, применялась Бейесова стратегия. При проблеме классификации в две категории точка x причисляется к категории

A , если

$$h_A \cdot l_A \cdot f_A(x) > h_B \cdot l_B \cdot f_B(x), \quad (2.6.2)$$

где

h_A и h_B — априорная вероятность нахождения принадлежащей к классу A или B точки,

l_A и l_B — величины, характеризующие потери, которые возникают при неправильной классификации относящихся к классам A и B точек.

(Вышеуказанное правило решения легко обобщается и для проблемы классификации на $M > 2$ категорий.)

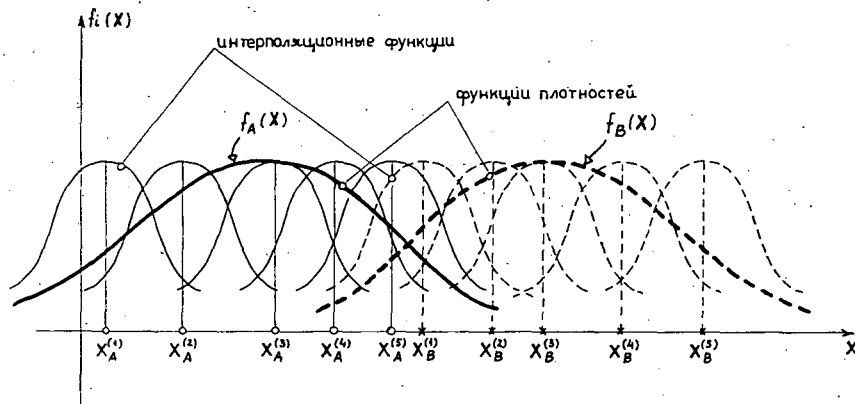


Рис. 1. Наложение интерполяционных функций

Однако, алгоритм, применяющий функции плотностей распределений вида (2.6.1) применим лишь при сравнительно небольшом количестве точек — представителей, т. к. в процессе обучения необходимо хранение всех точек обучающей последовательности, при этом классификация неизвестной точки требует пропорциональной длины обучающей последовательности вычислительной работы.

Этот недостаток устраняется разложением в ряд Тейлора выражения (2.6.1).

После разложения в ряд функция $f_A(x)$ разлагается — подробно описанным в [12] и [14] образом — на произведение независимой от точек обучающей последовательности положительной функции и на полином, поэтому для дальнейших классификаций достаточно применять этот полином. Коэффициенты полинома вычисляются из координат точек обучающей последовательности рекурсивным путем. В этом случае достаточно хранить вместо точек обучающей последовательности только коэффициенты полинома. Следовательно, потребность в памяти в процессе обучения постоянна.

Классификация неизвестных точек состоит из вычисления значения полинома подставляя координаты точки. Однако за вышеуказанные преимущества придется платить возникающей вследствие усечения ряда ошибкой. Опыт показывает, что при решении практических задач точность классификации не снижается значительно, по сравнению с методом классификации, применяющим функции плотностей вида (2.6.1).

Следует отметить, что применив ПДМ — в противоположность применения алгоритма потенциальных функций — полученная разделяющая функция не зависит от порядка следования точек обучающей последовательности. Эффективность алгоритма в значительной мере зависит от правильного выбора свободного параметра σ в (2.6.1) и участвующих в Бейсовом правиле решения коэффициентов потерь.

Программа, реализующая алгоритм для ЭВМ, состоит из рекурсивного вычисления коэффициентов полинома.

2.в. Алгоритм кусочно-линейного разделения

Известно, что линейным разделением решается лишь сравнительно узкий круг практических задач. Одним из возможных обобщений линейного разделения является метод кусочно — линейного разделения. Кусочно — линейную разделяющую поверхность можно создать соединением в сеть линейных решающих элементов (TLU: threshold logic unit).

Каждый решающий элемент (называется и пороговым элементом TLU), имеющий весовой вектор (вектор решения) w , классифицирует точку обучающей последовательности x по значению функции $\text{sign}(w, x)$.

Сеть, осуществляющая кусочно-линейное разделение (т. н. "committee machine") является простейшим типом многослойных сетей пороговых элементов.

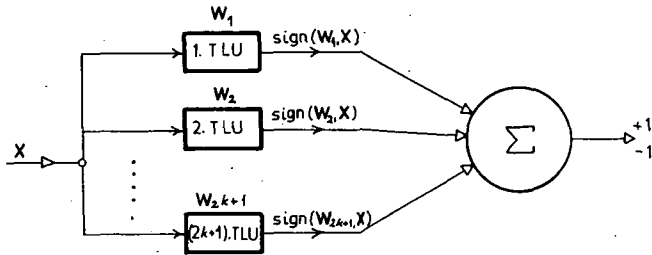


Рис. 2. "Committee machine"

Единственный находящийся во втором слое TLU считает «голоса» и решает на основании большинства (из этого следует, что в первом слое необходимо иметь нечетное число TLU).

В процессе обучения при неправильном решении устройства по заданному правилу изменяем векторы решений некоторых TLU. Например по описанному в [4] и использованному нами методу, к векторам решений нескольких определенных, неправильно классифицировавших элементов добавили произведение точки обучающей последовательности с заданным постоянным коэффициентом: $w'_j = w_j + C \cdot x$ где

w'_j — вектор решения TLU до учения,

w''_j — вектор решения TLU после учения,

C — постоянный коэффициент,

x — неправильно классифицированная точка обучающей последовательности.

Эта процедура повторяется до того, пока не добьемся правильной классификации точки x .

Для алгоритма кусочно-линейного разделения неизвестны теоремы о его сходимости. О сходимости обучения можно судить лишь на основании опытов.

На скорость обучения начальная настройка элементов имеет существенное влияние.

3. Метеорологическая модель задачи

Одним из типичных областей применения алгоритмов обучения является предсказывание погоды.

Применение в метеорологии методов распознавания образов и алгоритмов обучения излагается в [18], [20] и [21] в общем виде, а об успешном их использовании для конкретных метеорологических задач сообщается в [17] и [19]. В [17] сообщено об опытах, связанных с предсказыванием осадков, с помощью алгоритма кусочно-линейного разделения. В [18] описывается причисление к определенным погодным зонам отдельных территориальных частей СССР, с применением алгоритма самообучения, действующего по методу потенциальных функций. Судя по статьям, опубликованным в литературе, алгоритмы обучения дали согласные с ранее использованными методами, но в отдельных случаях превышающие их точность результаты. Однако, вопреки вышеуказанным результатам, в метеорологической практике машинные методы, применяющие алгоритмы обучения еще не получили распространение. Применение алгоритмов обучения испытано лишь в нескольких из многочисленных областей метеорологии.

В настоящей работе показывается решение задачи предсказывания конвективной активности, являющейся мерой грозовой деятельности атмосферы, с применением алгоритмов обучения.

В Венгрии в Центральном Метеорологическом Институте в последние годы была произведена интенсивная исследовательская работа в области метеорологических явлений, связанных с предсказыванием конвективной активности. Разработаны модель и метод для ЭВМ для объективного анализа условий окружения, благоприятствующих образованию конвективных процессов. Подробный отчет об этом находится в [22].

Индикаторы, разработанные в [22], составляли координаты точек обучающей последовательности, а последующая фактическая конвективная активность составляла учение.

Опыты, произведенные в процессе решения задачи, подробно описываются в [23].

Достигнутые результаты и получаемые из них выводы содержатся в [24]. Изложение результатов с точки зрения алгоритма ПДМ дается в [25].

В последующем дается описание в общих чертах существенных частей метеорологической модели, далее даются результаты, полученные применением разных алгоритмов, наконец, сравнение алгоритмов между собой.

В процессах атмосферных движений важную роль играют вертикальные потоки, которые появляются вследствие возмущений неустойчивых распределений воздушных масс.

Эти упорядоченные воздушные потоки называются конвекцией. Наиболее развитой их формой является конвекция дождевых кучевых облаков, что вертикально охватывает всю тропосферу, ее горизонтальные размеры порядка 10-ти км, а в линейной формации может достигать и несколько сотен км. Конвекция дождевых кучевых облаков всегда образуется при благоприятном содействии многочисленных факторов.

Принято считать самыми важными из этих факторов следующие:

- а) достаточное содержание водяного пара в нижних слоях воздуха;
- б) потенциальная неустойчивость гидростатического равновесия воздушного столба;
- в) существование механизма динамики, вызывающего выделения энергии неустойчивости.

В работе [22] описана процедура, применимая для ЭВМ, которая выявляет обстоятельства, определяющие формирование явлений конвекции дождевых кучевых облаков:

Для охарактеризования условий, при которых происходит образование конвективной активности, нами применены 12 параметров, выбранных таким образом, чтобы они с необходимой полнотой и весовым соотношением представляли вышеуказанные три фактора.

Для этого предикторы определены таким образом, чтобы каждый индикатор состоял из параметров одинаковой физической природы.

Определены всего 4 индикатора конвективной активности.

Индикатор I_1 состоит из гидростатических параметров. Среди его членов имеются: индекс устойчивости, индекс влажности, величина, характеризующая содержания влаги нижних уровней, и наконец параметр, характеризующий взаимность структур полей температуры и влажности.

Индикатор I_2 представляет собой развитие неустойчивости воздушного столба и состоит из разности геострофических адвекций температур на уровнях 500 и 850 миллибар.

Индикатор I_3 содержит такие параметры, которые указывают на существование механизмов, активизирующих скрытую неустойчивость.

Среди этих параметров находятся генез вихря скорости, вычисленная для уровня 850 мбар, а также функция генеза фронта накопления температуры и точки росы, далее геострофическая адвекция относительного вихря скорости.

Посредством индикатора I_4 попытались характеризовать одновременное присутствие развития неустойчивости и механизма активизации. Этот индикатор отмечает те области, внутри которых господствующими являются горизонтальная сходимости потока воздуха на нижних уровнях, а на средних уровнях горизонтальная расходимость потока воздуха; далее внизу наблюдаются адвекция теплого и влажного, а наверху адвекция сухого и холодного воздуха.

Иными словами I_4 отмечает те области, внутри которых наблюдается оптимальная комбинация развития неустойчивости воздушного столба, перестройка порядка вертикального распределения влаги, и триггер, возбуждающий конвективную циркуляцию.

Подводим итоги: целью настоящей работы является предсказывание величины конвективной активности. Из вышесказанного следует, что состояние атмосферы описывают 4 индикатора конвективной активности.

Площадь Западной и Средней Европы покрыли прямоугольной сетью, состоящей из 13×8 узловых точек сети. Таким образом получили $12 \times 7 = 84$ квадрата, состояние воздушного пространства внутри которых описывается индикаторами в 4-х точках сети (рис. 3).

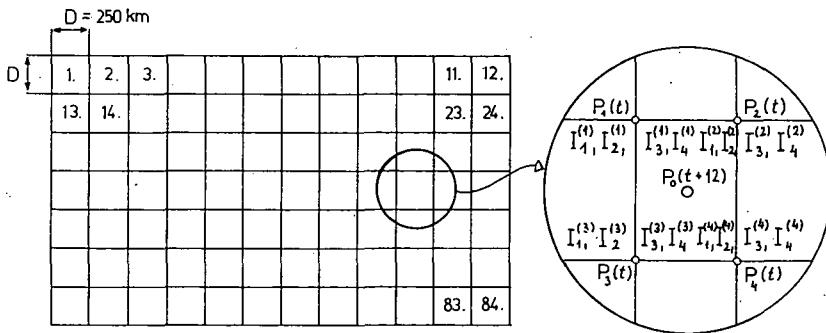


Рис. 3. Ортогональная сеть, покрывающая исследованную область

Каждый квадрат, длиной стороны 250 км, характеризуется 16 параметрами, т. е. при выработке существенных признаков задача отображается в 16-мерное Эвклидово пространство.

На основании этих 16 параметров следует определить появление или отсутствие конвективной активности в квадрате на по следующие 12 часов. Для опыта нами использованы индикаторы, вычисленные из данных, полученных высотными измерениями состояния воздуха 11-го июля 1968-го года в 12^{00} по Гринвичу. Параметр, описывающий фактическую конвективную активность в промежутке времени $12^{00} - 24^{00}$ по Гринвичу представляет собой учение.

Кроме 250-километровой сети, была применена и сеть с длиной стороны квадрата 125 км. Данные каждой сети были определены наложением прямоугольной сети на обработанные для сферической системы координат изолинии, и для каждой точки — по мере необходимости — были проведены интерполляции значений.

Величины параметра конвективной активности определялись следующим образом:

- а) Если не было конвективной активности, то $u = -1$,
- б) При редких ливнях $u = 1$,
- в) При многочисленных ливнях $u = 2$,
- г) При грозе, отмеченной какой-то станцией $u = 3$,
- д) При сильной грозовой деятельности $u = 4$.

4. Результаты, полученные применением программ и следствия, вытекающие из них

Каждый алгоритм из трех оказался пригодным для решения метеорологической задачи. Для проверки в случае алгоритма потенциальных функций, содержащего потенциальную функцию в виде полинома, и в случае алгоритма кусочно-линейного разделения была использована обучающая последовательность, что допустимо потому, что эти процедуры являются итеративными процедурами и автоматически не разделяют правильно на обучающую последовательность.

Однако для проверки алгоритма ПДМ была использована последовательность, отличающаяся от обучающей. Проверка производилась двумя различными методами. В случае первого метода данные находящиеся в нашем распоряжении были разделены на две части: первая часть была использована для обучения, а вторая для проверки (ПДМ/І). При втором методе обучение было проведено столько раз, сколько имелось точек обучающей последовательности. Из имеющихся точек была выделена одна, остальными было проведено обучение, а потом с помощью полученной разделяющей функции была произведена классификация оставшейся точки (ПДМ/ІІ).

Ниже показываются несколько характерных результатов.

В таблице H_1 и H_2 обозначают ошибки первого и второго родов. Ошибкой первого рода называется случай, когда алгоритм предсказывает ясную погоду, а в действительности появляется гроза.

Под точностью понимаем относительную частоту правильного предсказания.

В случае линейной разделяющей функции алгоритмы потенциальных функций дали следующие результаты: (таблица І).

Таблица І. Результаты алгоритмов потенциальных функций

Алгоритм	Перфоленга данных	Точки проверки (тест)	H_1	H_2	$H=H_1+H_2$	Точность (%)
Тета	«250 km»	$84 = 33_{гр.} + 51$	9	6	15	82.2
Гамма	«250 km»	$84 = 33_{гр.} + 51$	14	10	24	71.5
Тета	«125 km»	$336 = 92_{гр.} + 244$	20	30	50	85.2
Гамма	«125 km»	$336 = 92_{гр.} + 244$	36	21	57	83.1

Из таблицы видно, что тета-алгоритм, у которого величина коррекции не зависит от положения точки внутри обучающей последовательности, дал лучшие результаты чем гамма-алгоритм, где величина коррекции убывает в процессе обучения. При применении разделяющих функций более высокого порядка чем линейные точность не увеличивалась, что наводит на мысль о том, что положение не таково будто бы два множества в 16-и мерном пространстве, соответствующие состояниям «гроза» и «не гроза», являются непересекающимися, только линия раздела очень сложна, а наоборот, эти два множества пересекаются (см. рис. 4).

В случае итеративного алгоритма, выработанная разделяющая функция подвергается влиянию порядка следования точек обучающей последовательности. Поэтому, у алгоритмов потенциальных функций нами было проведено и испытание с целью определения того, как влияет порядок следования точек обучающей последовательности на точность разделения.

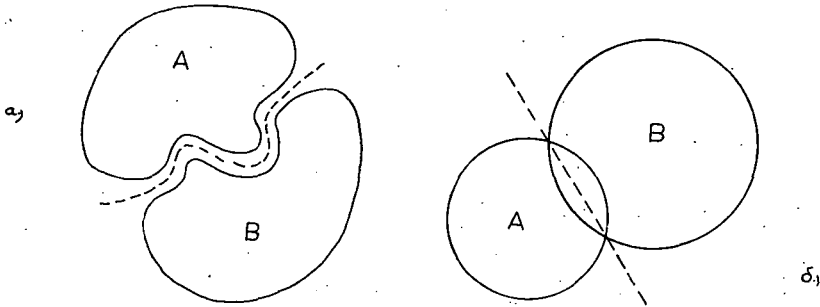


Рис. 4. Характер разделяемых множеств

Для этой цели были изготовлены несколько «смешанных» последовательностей данных. «Смешанная» последовательность получена следующим образом: последовательность точек, полученных для 250-километровой сети смешали относительно их «естественного» порядка следования (см. рис. 6). Результаты, полученные такими «смешанными» данными, показаны в таблице II и получены с применением линейного тета-алгоритма.

Таблица II. Результаты, полученные применив несколько «смешанных» последовательностей данных

Перфолента данных	Число однородных отрезков	N_1	N_2	N	Точность (%)
«естественная»	15	9	6	15	82.2
«случайная»	35	8	4	12	85.7
«равномерная/1»	67	6	6	12	85.7
«равномерная/2»	67	5	5	10	88.1

В первой строке таблицы содержатся результаты последовательности данных с «естественной» очередностью. Так как грозовые зоны образуются обычно на связных областях, поэтому здесь получено относительно небольшое число однородных отрезков (внутри однородного отрезка находятся либо одни «грозовые» точки, либо одни точки «ясной погоды»). Во второй строке показан результат такой последовательности данных, где порядок следования точек установлен с помощью генератора случайных чисел. Здесь число однородных отрезков больше.

В третьей и четвертой строках показаны результаты таких последовательностей данных, где точки обучающей последовательности поставлены в поря-

док следования, при котором старались придать распределению «грозових» точек как можно большую равномерность.

Здесь число однородных отрезков максимальное. Из опытов видно, что имеется тенденция к повышению точности предсказания с увеличением числа однородных отрезков.

Характерные результаты алгоритма ПДМ содержатся в таблице III.

Таблица III. Результаты алгоритма ПДМ

Метод испытания	Перфолента данных	Серия тестов	H_1	H_2	H	Точность (%)
ПДМ/І	«250 km»	42 = 22гр. + 20	4	1	5	88.1
ПДМ/ІІ	«250 km»	84 = 33гр. + 51	9	8	17	80.0
ПДМ/ІІІ	«125 km»	168 = 35гр. + 133	4	16	20	88.1

Алгоритм ПДМ надежно работает. Разделяющая функция, полученная при помощи обучающей последовательности, очень хорошо классифицирует также и серии тестов, отличающихся от обучающей последовательности.

Результаты, полученные с применением алгоритма кусочно-линейного разделения, приведены в таблице IV. Опыты, проведенные нами, показали, что в общем случае алгоритм кусочно-линейного разделения существенно сходится со скоростью, существенно меньшей, чем алгоритм потенциальных функций. Результаты, полученные применив исходную обучающую последовательность, оказались довольно слабыми (1). Поэтому нами была увеличена длина обучающей последовательности таким образом, что произвели обучения с простым 5-, 10-кратным повторением исходной обучающей последовательности. Удлинив таким образом обучающую последовательность, мы получили существенно улучшенные результаты (2 и 3).

Таблица IV. Результаты алгоритма кусочно-линейного разделения

Длина обучающей последовательности	H_1	H_2	H	Точность (%)
1-кратная	24	2	26	69.1
5-кратная	15	0	15	82.2
10-кратная	2	4	6	92.8

Результаты получены с применением 3 TLU и при значении корректирующего множителя $C=1$. Исходная обучающая последовательность состояла из 84 точек 250-километровой сети. Увеличение числа TLU не дало улучшение точности классификации.

Наконец, нами исследовалось и то, есть ли возможность уменьшать число мер. Результаты, приведенные в таблице V, показали, что задание с данными 16-мерного пространства является в значительной мере излишним.

В первой строке обозначение R3/1, 3, 4 означает, что из 4-х индикаторов для каждой точки один остался неиспользованным, и так применялись I_1 , I_3

Таблица V. Результаты данных с уменьшенным числом мер

№	Вид сокращения	Число мер	H_1	H_2	H	Точность (%)
1	R3/1, 3, 4	12	7	6	13	84.5
2	R2/1, 4	8	7	7	14	83.1
3	R1/4	4	4	15	19	77.4
4	R4/среднее	4	11	7	18	78.5
5	R2/1, 4/среднее	2	7	8	15	82.2
6	R2/1, 4/среднее	2	4	2	6	85.7

и I_4 . Таким образом проблема стала сокращенной в 12-мерное пространство. Во второй строке R2/1, 4 означает, что применены лишь индикаторы I_1 и I_4 , тогда проблема станет 8-мерной. Из результатов видно, что в 12-мерном и даже в 8-мерном пространстве разделение осуществимо без уменьшения точности. Далее (3) видно, что при использовании лишь индикатора I_4 , тем самым сократив проблему в 4-мерное пространство, точность не уменьшается значительно. Проблему можно сократить в 4-мерную и другим способом: образованием среднего из значений индикаторов, принятых ими в четырех углах квадрата. В этом случае нами достигнута приблизительно такая же точность разделения, что и при другой четырехмерной проблеме.

Особенный интерес представляют собой результаты, находящиеся в 5-ой и 6-ой строках таблицы V. В этих случаях проблема уже двухмерная, что достигнуто таким образом, что в последовательности вышеописанным способом вычислены средние значения. Применив эту двухмерную обучающую последовательность, алгоритм обучения классифицировал с такой же точностью, что и в случае исходного 16-мерного пространства данных! А программа ПДМ/И с такой же перфолентой данных, когда для обучения применялись первые 42 точки, а для проверки остальные, дала очень хорошую точность классификации (приблизительно 86%), что также очень мало отличалась от 88%-ного результата, полученного применив 16-мерную обучающую последовательность при подобных обстоятельствах. Ясно видно на основании опыта, полученного в связи с сокращением данных, что применение алгоритмов имеет значение также с точки зрения усовершенствования метеорологической модели, так как с их помощью можно судить об относительной важности каждого индикатора, с точки зрения точности предсказания. Можно установить первостепенную важность индикатора I_4 , ибо достигаемая на основании одного I_4 точность предсказания мало отстает от точности получаемой при одновременном применении всех четырех индикаторов.

Вторым по важности можно отметить I_1 , так как применив I_1 и I_4 точность достигает значения, которое имелось при применении всех четырех. Менее важны индикаторы I_2 и I_3 , т. к. их отсутствие не имеет заметное влияние на точность классификации.

Результаты, находящиеся в табл. V, получены линейным тета-алгоритмом потенциальных функций, применив данные 250-километровой сети, за исключением 6-ой строки, полученной алгоритмом ПДМ/1.

Наконец, была исследована и причина появления ошибочных решений алгоритмов. Для этого сравнивали результаты девяти различных алгоритмов — разных вариантов показанных ранее алгоритмов — и прицели решение на основании большинства. Таким путем получены приведенные в таблице VI результаты.

Результаты, полученные применив вышеуказанные 9 алгоритмов, изображены на рис. 5. На рис. 5 изображена квадратическая сеть натянутая на карту Европы с длиной стороны квадрата 250 км. Цифра, находящаяся в левом верхнем углу квадрата, является его порядковым номером. Отметим для ориентировки, что Лондон находится в 40-ом, а Будапешт в 58-ом квадратах.

Цифры, находящиеся в правом, нижнем углу квадратов, показывают, сколько из названных 9-ти алгоритмов дали неправильное решение. Буквы внутри квадрата отмечают, какие именно алгоритмы давали неправильное решение согласно следующему коду:

— от а) до г) буквы отмечают результаты линейного тета-алгоритма первоначально принятого вида при применении обучающей последовательности с первоначальным порядком следования точек (g) и обучающими последовательностями различных видов «смешанных» порядков следования точек.

— h) относится к результатам, приведенным в строке 2 таблицы IV, полученным с применением алгоритма кусочно-линейного разделения.

— i) отмечает результаты приведенные в строке 2 таблицы III алгоритма ПДМ II.

Штрихованные квадраты отмечают те территории, на которых в действительности была гроза, а обрамленные жирной линией квадраты соответствуют тем территориям, где *большинство алгоритмов* дало *неверное решение*.

Общее количество таких квадратов 11, следовательно решение, принятое на основании решения большинства алгоритмов, является правильным в 87%-ах.

Пропорции решений алгоритмов:

9:0	в 45 случаях	} Правильное решение: в 74 случаях (87%)
8:1	в 14 случаях	
7:2	в 7 случаях	
6:3	в 4 случаях	
5:4	в 3 случаях	

4:5	в 2 случаях	} Неправильное решение: в 11 случаях (13%)
3:6	в 2 случаях	
2:7	в 1 случае	
1:8	в 4 случаях	
0:9	в 2 случаях	

1. <i>b, c, d, e, f, g, h, i</i> 8	2. <i>b, c, d, e, f, i</i> 6	3. 0	4. 0	5. 0	6. <i>e</i> 1	7. 0	8. 0	9. 0	10. 0	11. 0	12. 0
13. 0	14. <i>c, e, f</i> 3	15. <i>c</i> 1	16. <i>b, c, d, e, f, g, h, i</i> 8	17. <i>c, i</i> 2	18. <i>b, c, d, g</i> 4	19. <i>b, c, d, g, h, i</i> 6	20. <i>b, d, g, h, i</i> 5	21. <i>a, c, e, f</i> 4	22. <i>a, b, c, d, e, f, g, h, i</i> 9	23. <i>f</i> 1	24. 0
25. 0	26. <i>f</i> 1	27. <i>a, b, c, d, e, f, g, h, i</i> 9	28. /	29. 0	30. <i>l, h</i> 2	31. /	32. /	33. <i>a</i> 1	34. <i>a, e, f</i> 3	35. <i>b, c, g, h, i</i> 5	36. <i>f</i> 1
37. 0	38. 0	39. 0	40. 0	41. <i>a, c, d, e, g, h, i</i> 7	42. <i>h</i> 1	43. /	44. /	45. <i>a</i> 1	46. /	47. <i>a</i> 1	48. <i>a, b, c, d, f, g, h, i</i> 8
49. 0	50. 0	51. 0	52. 0	53. <i>a, f, h</i> 3	54. /	55. <i>a, f</i> 2	56. <i>a, f</i> 2	57. /	58. /	59. <i>d, g</i> 2	60. 0
61. 0	62. <i>f</i> 0	63. <i>f</i> 1	64. <i>i</i> 1	65. <i>a, b, c, e, f, g, h, i</i> 8	66. /	67. <i>a</i> 1	68. <i>a</i> 1	69. <i>g, i</i> 2	70. /	71. 0	72. 0
73. 0	74. 0	75. /	76. <i>a</i> 1	77. <i>a, g</i> 2	78. <i>a, b, i</i> 3	79. <i>a, d, g, i</i> 4	80. /	81. /	82. /	83. 0	84. 0

Рис. 5. Территориальное размещение неправильных решений

Наблюдения показали, что ошибки, сделанные таким образом, были совершены в тех областях, которые разместились на границе грозовой зоны или происходили в тех областях, где данные измерения не являлись достоверными. Следовательно, появившиеся ошибки главным образом происходят из несовершенства метеорологической модели, а также из неточности измеренных данных. На это указывает и то обстоятельство, что тогда, когда исключили из обучающей последовательности те 11 точек, для которых решения большинства алгоритмов дали неправильный результат,

Таблица VI. Результат решений большинства девяти алгоритмов

N_1	N_2	N	Точность (%)
5	6	11	87

то применив обучающую последовательность из 73-х остальных точек, линейный тета-алгоритм потенциальных функций классифицировал с очень высокой точностью (см. таблицу VII) при однократном пропуске обучающей последовательности с точностью

Таблица VII. Рост точности при отбрасывании не надежных точек обучающей последовательности

Кратность пропусков	N_1	N_2	N	Точность (%)
1	3	1	4	94.6
3	2	0	2	97.4

94,6%, а при трехкратном пропуске — с точностью 97,4 %.

Подытожив все вышеизложенное, можно сказать, что в случае изложенной выше метеорологической модели конвективная активность хорошо предсказывается с помощью показанных алгоритмов обуче-

ния. Надежность предсказания превышает надежности синоптической практики; а кроме этого: главным достоинством процедур-применяющих алгоритм обучения по сравнению с синоптической практикой является то, что они дают объективный метод для предсказания.

Можем рассчитывать на дальнейшее увеличение точности предсказания путем увеличения длины обучающей последовательности.

В ближайшем будущем мы расширим круг исследований на предсказывания распределения во времени, — до сих пор нами исследовалось лишь пространственное распределение — а потом также и на предсказывание совместного пространственно-временного распределения. Из этих исследований рассчитываем получить такой опыт, который позволит усовершенствовать метеорологическую модель.

Application of learning algorithms in meteorology for the prediction of the convective activity

In this paper the application of some supervised learning algorithms in connection with meteorological prediction is described. The meteorological task, solved in cooperation of the Central Institute of Forecasting and the Research Institute for Telecommunication, related to the prediction of convective activity, or more popularly, to the prediction of rainstorms.

First a brief survey of the general theory of supervised learning is given, then the algorithms used in our experiments are discussed, and the computer programs realizing them are presented. For solving the above mentioned task three algorithms were used: the method of potential functions, introduced by Aizerman, Braverman, Rozonoer and others [6], the polynomial discriminant method of Specht [12] and the piecewise-linear separator of Nilsson [3] called committee machine. In the fol-

lowing the most important details of the meteorological model, thoroughly discussed in [22], is given. Finally the results and conclusions are discussed.

On the basis of the results it can be stated that the learning algorithms used in our experiments are able to predict the convective activity with high reliability, which exceeds the usual standards of more traditional techniques in meteorology.

* CENTRAL INSTITUTE OF FORECASTING,
BUDAPEST, HUNGARY
II. KITAIBEL PÁL U. 1.

** RESEARCH INSTITUTE FOR TELECOMMUNICATION,
BUDAPEST, HUNGARY
II. GÁBOR ÁRON U. 65.

Литература

- [1] Васильев, В. И., Распознающие системы, *Наукова Думка*, Киев, 1969, стр. 292.
- [2] Цыпкин, Я. З., Адаптация, обучение и самообучение в автоматических системах, *Автоматика и Телемеханика*, т. 27, 1966, стр. 23—61.
- [3] NILSSON, N. J., *Learning machines*, New York, McGraw Hill, 1965.
- [4] CSIBI, S. & O. GULYÁS, Задачи по распознаванию образов (на венгерском языке), ТКИ, *Szemináriumi Közlemények*, 1969.
- [5] Башкиров, О. А., Э. М. Браверман, И. Б. Мучник, Алгоритмы обучения распознаванию зрительных образов, основанные на использовании потенциальных функций, *Автоматика и Телемеханика*, т. 25, 1964, стр. 692—695.
- [6] Айзерман, М. А., Э. М. Браверман, Л. И. Розоноер, Теоретические основы потенциальных функций в задаче об обучении автоматов разделению входных ситуаций на классы, *Автоматика и Телемеханика*, т. 25, 1964, стр. 917—936.
- [7] Браверман, Э. М., О методе потенциальных функций, *Автоматика и Телемеханика*, т. 27, 1966, стр. 2205—2213.
- [8] GULYÁS, O., Метод потенциальных функций (на венгерском языке), ТКИ, *Szemináriumi Közlemények*, 1969.
- [9] GULYÁS, O., On extended potential function type learning algorithms and their convergence rate, *Problems of Control and Information Theory*, v. 1, 1972, pp. 51—64.
- [10] HOFFER, A., Распознавание образов с помощью алгоритма обучения, О выборе потенциальной функции, Дипломная работа (на венгерском языке), ELTE, ТТК, Кафедра теории вероятностей, 1970.
- [11] MOLNÁR, L. & O. GULYÁS, Программа обучения по методу потенциальных функций (на венгерском языке), ТКИ, *Szemináriumi Közlemények*, 1969.
- [12] SPECHT, D. F., Generation of polynomial discriminant functions for pattern recognition, *IEEE Trans.*, v. EC—16, 1967, pp. 308—319.
- [13] SPECHT, D. F., Vectorcardiographic diagnosis using the polynomial discriminant method of pattern recognition, *IEEE Trans.*, v. BME—14, 1967, pp. 90—95.
- [14] MEISEL, W. S., Potential functions in mathematical pattern recognition, *IEEE Trans.*, v. C—18, 1969, pp. 911—918.
- [15] MOLNÁR, L., Метод распознавания образов полиномиальной дискриминацией (на венгерском языке), ТКИ, *Szemináriumi Közlemények*, 1970.
- [16] ESZE, T. & J. NÉMETH, Линейное и кусочно-линейное разделения (на венгерском языке), ТКИ, *Szemináriumi Közlemények*, 1969.
- [17] WIDROW, B. & F. W. SMITH, *Pattern recognizing control systems*, Washington, 1964, pp. 288—317.
- [18] Сонечкин, Д. М., Математическая теория классификации и ее применение в метеорологии, *Метеорология и гидрология*, 1969, стр. 24—34.
- [19] Сонечкин, Д. М., Об объективной классификации метеорологических явлений и ситуаций с помощью ЭВМ, *Метеорология и гидрология*, 1968, стр. 12—21.
- [20] Баргов, Н. А., О классификации синоптических процессов, *Метеорология и гидрология*, 1969, стр. 3—12.
- [21] Груза, Г. Б., Прогноз погоды и задача «распознавание образов» в кибернетике, *Метеорология и гидрология*, 1968, стр. 13—21.
- [22] SZALAY, G. & G. GÖTZ, Метод для объективного анализа условий окружения конвективных процессов (на венгерском языке), *Időjárás*, v. 75, 1971, pp. 90—102.

- [23] SZALAY, G., L. MOLNÁR, O. GULYÁS, Применение алгоритмов обучения для метеорологического прогнозирования (на венгерском языке), *Междуинститутский доклад*, KEI—TKI, 1970.
- [24] SZALAY, G., L. MOLNÁR, O. GULYÁS, Использование алгоритмов обучения для метеорологического прогнозирования (на венгерском языке), Доклад на *VI-ой Венгерской конференции по автоматизации*, Будапешт, 1970.
- [25] MOLNÁR, L., The polynomial discriminant method of pattern recognition and its use in meteorology (paper presented at *The VIth Yugoslav International Symposium on Information Processing*, Bled, 1970).

(Поступило 8-ого марта 1971 г.)