

УДК 004.032.26

Т. Р. Жангиров, А. С. Перков, С. А. Иванова, А. А. Лисс
Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Н. Ю. Григорьева
СПб НИЦЭБ РАН

Л. В. Чистякова
РЦ Культивирование микроорганизмов НП СПбГУ

Сравнение эффективности решения задачи классификации методами линейного дискриминантного анализа и искусственных нейронных сетей

Приводится сравнение двух методов классификации: линейного дискриминантного анализа и искусственных нейронных сетей прямого распространения. Сравнение проводится на примере задачи классификации цианобактерий по серии спектров собственной флуоресценции отдельных клеток. Продемонстрированы способы визуализации и анализа результатов классификации, включая новые способы, основанные на энтропии результатов классификации. Определена связь между результатами линейного дискриминантного анализа и классифицирующих нейронных сетей. Показаны ключевые моменты, на основе которых можно выявить некорректно обученную модель нейронной сети, а также предложен новый способ анализа качества обобщения. Представленные в работе методики сравнения результатов линейного дискриминантного анализа и нейронных сетей применимы для сравнения нейронных сетей с другими алгоритмами машинного обучения и статистическими классификаторами. Предложенная методика позволяет анализировать процесс обучения нейронной сети с целью оперативного выявления критичных отклонений в ее работе и возможного алогичного поведения в дальнейшем.

Классификация, линейный дискриминантный анализ, нейронные сети, статистический анализ, градиентный спуск

Задача классификации – одна из задач обработки данных, заключающаяся в сопоставлении наблюдения, характеризующегося набором признаков, и одного из заранее известных целевых классов [1]. Основная проблема решения данной задачи заключается в построении корректной модели с минимально возможной ошибкой классификации. Существует большой набор методов классификации (методов построения модели). Выбор метода классификации производится на основе предметной области и характера данных. Классификация наблюдений наиболее востребована в областях, где имеются прикладные задачи, связанные с проведением экспериментов и большим объемом данных. Примерами таких областей могут быть медицина, экология, микробиология, геология, электроника и т. д. В данной статье проводится сравнение результатов линейного дискриминантного анализа (ЛДА) [2] и классифицирующих нейронных сетей [3], [4].

Линейный дискриминантный анализ – метод классификации, основанный на анализе статистических характеристик выборки наблюдений. Вследствие этого основным достоинством ЛДА является то, что для одной и той же выборки всегда будет получаться одна и та же модель классификатора. Тем самым за счет отсутствия случайной составляющей в методе его можно использовать как эталон для сравнения с методами классификации, имеющими случайную составляющую, например, с нейронными сетями. С другой стороны, недостатком ЛДА можно считать то, что не всегда есть возможность провести анализ, если между признаками наблюдается сильная корреляция [5].

Искусственная нейронная сеть (ИНС) – одна из моделей машинного обучения, позволяющая решать задачи эвристическим путем. С помощью нейронных сетей можно решать целый ряд задач,

одной из которых является классификация. Благодаря гибкости и адаптивности модели, нейронные сети нашли широкое применение в разных областях, особенно там, где невозможно использовать стандартные методы обработки данных [6]–[8]. Основными достоинствами применения классификационных нейронных сетей являются, во-первых, возможность создавать модели с высокой степенью нелинейности и, во-вторых, возможность работы с выборками, которые не всегда можно качественно статистически проанализировать. Одним из существенных недостатков, который связан с природой нейронных сетей, является случайность построенной модели. На одной и той же выборке можно получить большое количество различных моделей, которые могут давать одинаково хороший результат, что ведет к необходимости оценке таких моделей и выбору наилучшей из них. В данной работе продемонстрированы некоторые способы оценки результатов работы классифицирующих ИНС, в том числе основанные на сравнении с ЛДА.

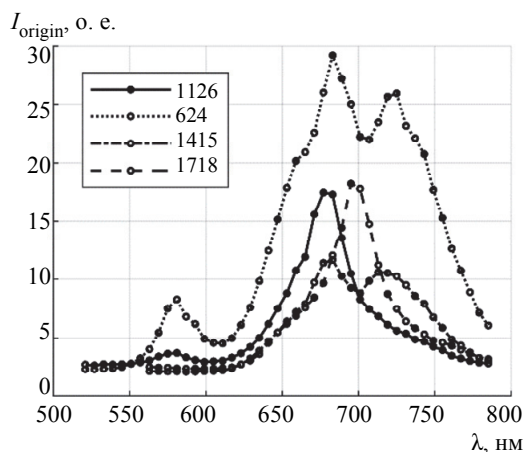
Сравнение ИНС и ЛДА проводилось на основе решения прикладной задачи микробиологии, а именно, классификации 16 штаммов цианобактерий на основе серии спектров собственной флуоресценции. Основные трудности при решении данной задачи состояли в малом объеме исходной выборки, а также в сложном характере экспериментальных данных.

Исходные данные. В данной работе для классификации цианобактерий используются спектры собственной флуоресценции, снятые с отдельных клеток *in vivo* на конфокальном лазерном сканирующем микроскопе. Эта методика была недавно разработана одним из авторов данной статьи. В отличие от исследования интегральных характеристик суспензии клеток, исследование серий спектров с отдельных клеток позволяет уменьшить влияние различных побочных факторов, влияющих на форму кривой спектра собственной флуоресценции и осложняющих обработку экспериментальных данных. Кроме того, исследование живых клеток (*in vivo*) дает более полную и однозначную информацию о родовой принадлежности отдельных штаммов цианобактерий и способствует более точной дифференциации штаммов. В рамках данной методики каждое наблюдение характеризуется серией из 8 спектров собственной флуоресценции отдельных клеток, полученных при 8 различных длин волн

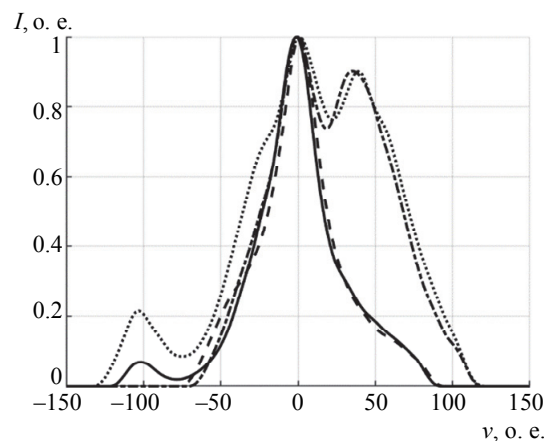
возбуждающего лазера (405, 458, 476, 488, 496, 514, 543, 633 нм). В работе каждый класс обозначается коллекционным номером из базы коллекции CALU Ресурсного центра «Культивирование микроорганизмов» НП СПбГУ.

Основной проблемой при обработке исходных данных является малое количество экспериментальных точек в спектре собственной флуоресценции. Это связано с тем, что под воздействием интенсивного лазерного излучения клетка может погибнуть, и поэтому необходимо ограничивать время воздействия лазера. Другая проблема связана с тем, что абсолютные значения спектров флуоресценции зависят от конкретных настроек конфокального микроскопа, поэтому необходимо привести все серии спектров к одному универсальному виду. В рамках данного исследования был разработан алгоритм унификации спектров флуоресценции, который включает в себя интерполяцию кубическим сплайном [9], нормализацию спектров, выравнивание их по оси частот и центрирование по длине волны 682 нм, сглаживание прямоугольной оконной функцией и функцией Ханна [10], а также экстраполяцию в заданном диапазоне длин волн. Такая предварительная обработка исходных данных позволяет восстановить пропущенные значения спектра и осцилляции спектра. На рис. 1, а, б представлены исходные спектры, снятые при длине волны возбуждающего лазера 458 нм для четырех рассматриваемых классов с номерами 1126, 624, 1415, 1718 до (а) и после (б) процедуры обработки. Далее из сглаженных и нормированных спектров извлекаются различные классификационные признаки. До обработки каждый спектр отображает зависимость исходной интенсивности излучения (I_{origin}) от длины волны излучения (λ), а после обработки – нормированное на максимум значение интенсивности излучения (I) от откалиброванных значений (ν).

Классификационные признаки. После предварительной обработки каждый спектр характеризуется 600 точками, а каждое наблюдение 8 спектрами собственной флуоресценции. Проводить на таком объеме классификацию вычислительно сложно и неэффективно, так как частично форма спектров для цианобактерий, принадлежащих разным штаммам, имеет аналогичный вид. Поэтому необходимо выделить из каждой серии полезную информацию, которая достаточна для описания всех особенностей спектра и может



а



б

Рис. 1

служить характеристикой данного класса наблюдений. В предыдущих работах часть классификационных признаков уже была описана [11], а именно, асимметрия (А) и эксцесс (Э) и относительная интенсивность флуоресценции отдельных пигментов (ОИФ). Однако в данном исследовании будут использоваться другие четыре области спектра для параметра ОИФ: [573 586], [649 661], [674 689], [711 727] (рис. 2).

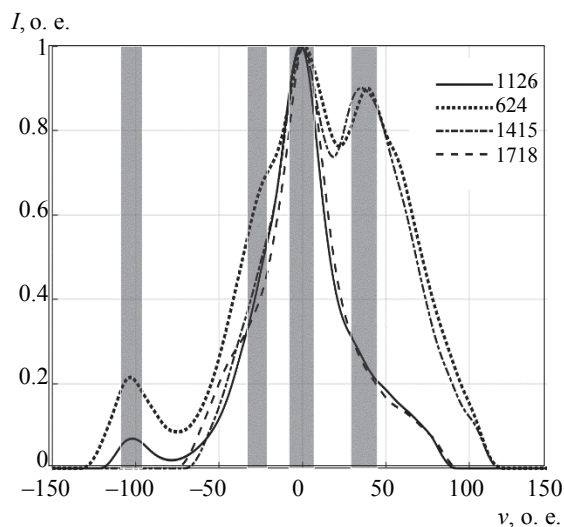


Рис. 2

Помимо описанных ранее классификационных признаков, в данном исследовании будет использоваться признак, основанный на Фурье-трансформанте [12] первой производной спектра собственной флуоресценции. Так как спектр представлен дискретно, то для нахождения производной используем формулу двухсторонней разности:

$$I'_i = (I_{i-1} - I_{i+1})/2,$$

где для определения производной в точке i (I'_i) применяются значения интенсивностей в предыдущей I_{i-1} и последующей I_{i+1} точках.

На рис. 3 представлены графики производных для соответствующих исходных спектров (рис. 1 и 2). Как следует из представленных кривых, на всех графиках наблюдаются высокочастотные колебания, вызванные наличием шумов в соответствующих спектрах. В связи с этим при анализе Фурье-трансформанты необходимо анализировать только низкочастотные составляющие.

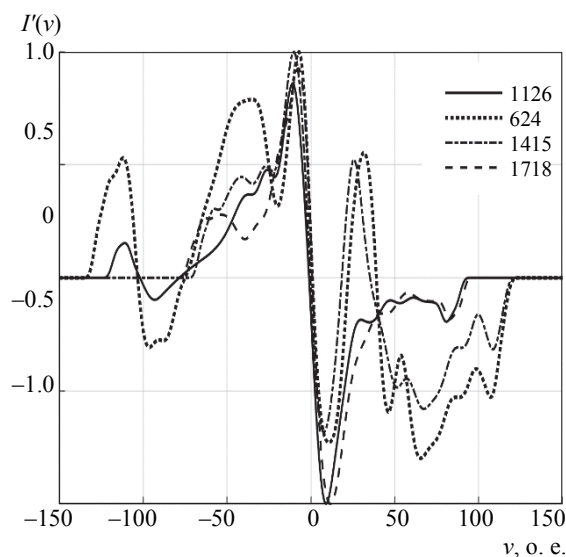


Рис. 3

Для расчета Фурье-трансформанты применялось быстрое преобразование Фурье (БПФ) с частотой дискретизации 1/2048.

$$F_k = \sum_{j=-1024}^{1024} I'_j e^{-\pi i(j-1)(k-1)/1024},$$

где F_k – значение Фурье-амплитуды в соответствующих точках k , I'_j – амплитуды первой производной спектра собственной флуоресценции в соответствующих точках j , k – нумерует точки Фурье-трансформанты, j – нумерует точки первой производной спектра собственной флуоресценции.

Таким образом, можно рассчитать амплитудную и фазовую характеристики производной исходного сигнала. В силу того, что производная спектра флуоресценции не является периодическим сигналом, ее фазовая характеристика не несет полезной информации. В данном случае интерес представляет только амплитудная характеристика, которая рассчитывается как модуль Фурье-трансформанты.

На рис. 4 представлена низкочастотная часть амплитудной характеристики для рассматриваемых производных исходных спектров (рис. 3). Для удобства анализа амплитудная характеристика нормирована по максимуму значения и рассматриваются только первые 50 отсчетов, соответствующие низкочастотной составляющей.

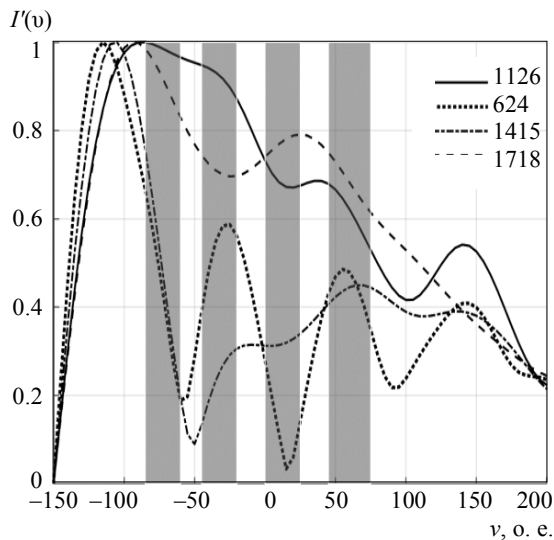


Рис. 4

Для выделения классификационных признаков на рис. 4 определены четыре характерные области, в которых рассчитано среднее значение амплитуды и процентный вклад каждой области в общую сумму Φ_α , аналогично параметру ОИФ:

$$\Phi_\alpha = \alpha_i / \sum_{j=1}^N \alpha_j,$$

где α_j – среднее значение амплитудной характеристики в области j ; α_i – среднее значение амплитудной характеристики в области i .

Таким образом, в качестве классификационных признаков в дальнейшем анализе будут выступать: асимметрия (А), эксцесс (Э) и относительная интенсивность флуоресценции отдельных пигментов (ОИФ), а также процентный вклад четырех областей амплитудной характеристики (ОАХ).

Линейный дискриминантный анализ. Линейный дискриминантный анализ позволяет строить линейные дискриминирующие функции, на основе которых можно классифицировать наблюдения:

$$f_y(\mathbf{x}) = \mu_y \Sigma_W^{-1} \mathbf{x}^T - 0.5 \mu_y \Sigma_W^{-1} \mu_y^T + \ln(1/Y),$$

где μ_y – вектор средних значений признаков класса y ; Σ_W – матрица внутриклассовых ковариаций; Y – количество целевых классов; $\mathbf{x}$ – вектор переменных функции f_y , представляющих набор признаков наблюдения.

Количество разделяющих функций равняется количеству целевых классов Y . Целевой класс, к которому относятся наблюдения, определяется как $\arg \max_y f_y(\mathbf{x})$, где $f_y(\mathbf{x})$ – дискриминирующая функция, соответствующая классу y .

Одним из преимуществ ЛДА является то, что для каждого класса можно оценить значимость признаков оценивания значения коэффициентов соответствующей дискриминантной функции.

Со стандартным ЛДА тесно связан линейный дискриминант Фишера [2], который является разновидностью канонического корреляционного анализа, позволяющий оценить качество набора признаков, а также аналогично методу главных компонент уменьшить размерность пространства признаков. Трансформирующая матрица T составляется из собственных векторов-столбцов матрицы $\Sigma_W^{-1} \Sigma_B$ (Σ_B – матрица межклассовых ковариаций), причем в матрицу входят только уникальные векторы, соответствующие положительным вещественным собственным числам. Для удобства столбцы матрицы T располагаются в порядке уменьшения соответствующих собственных чисел. Таким образом, первая координата в новом пространстве отвечает за максимальное разделение между классами.

Искусственные нейронные сети. В данной работе для решения задачи классификации выбрана многослойная нейронная сеть прямого распространения [13]. В такой сети распространение сигналов происходит от входного слоя к выходному, остальные слои называются скрытыми. В рамках задачи классификации значения сигналов на входном слое соответствуют значениям признаков для отдельного наблюдения, а значения на выходном слое отображают принадлежность к

определенному классу. Активационная функция, используемая в нейронной сети, определяет, как состояние нейрона преобразуется в выходной сигнал. В данной работе в качестве таких функций используется гиперболический тангенс на скрытых слоях и функция Softmax на выходном слое. Функция Softmax обычно используется на выходном слое при решении задачи классификации, так как носит вероятностный характер и рассчитывается на основе состояний всего слоя. Значения для одного выходного нейрона находятся в диапазоне $[0, 1]$, причем сумма выходов для всего слоя равна 1. Исходя из этого, появляется возможность работать с нечеткой логикой, а также проверять наблюдения, которые могут не относиться ни к одному из известных нейронной сети классов. При хорошо обученной сети для таких наблюдений вероятностное распределение выхода сети будет показывать, к каким классам они относятся с наибольшей вероятностью.

Так как практически невозможно сразу правильно инициализировать веса для корректной работы нейронной сети, ее необходимо обучить. Обучение нейронной сети – это подбор весовых коэффициентов с целью минимизации ошибки в рамках задачи классификации: минимизации неправильно классифицируемых наблюдений. В данной работе для обучения ИНС используется метод обратного распространения ошибки [14], основная идея которого заключается в расчете градиента ошибки сети относительно каждого веса и в изменении каждого веса в соответствии с градиентом. В задачах классификации, где выход сети представляет вероятностное распределение, ошибка рассчитывается по формуле кросс-энтропии:

$$E(Y, C) = -\sum_i C_i \ln Y_i,$$

где Y – значения, полученные на выходе ИНС; C – значения, ожидаемые на выходе ИНС.

После расчета градиента ошибки относительно весов производится коррекция весов. В данной работе используется формула с поправкой Нестерова [15]:

$$\Delta w[n] = \eta(\nabla E + \rho w[n-1]) + \mu \Delta w[n-1],$$

где n – номер итерации обучения; η – скорость обучения; ρ – коэффициент регуляризации; μ – коэффициент момента.

В настоящей статье не приводится алгоритм нахождения гиперпараметров нейронной сети, приводящих к наилучшему результату, который,

по сути, заключается в переборе различных комбинаций гиперпараметров. Отметим, что в ходе проведенного исследования оказалось, что оптимальной нейронной сетью для решения поставленной задачи является сеть с одним скрытым слоем, размер которого составляет 31 нейрон, и использованием нейронов сдвига. Размер входного и выходного слоя равняется количеству классификационных признаков и целевых классов соответственно. Для алгоритма обучения оптимальными значениями гиперпараметров являются: $n = 2500$, $\eta = 0.01$, $\rho = 0.0001$, $\mu = 0.0001$.

Результаты и обсуждения. Основным показателем качества классификации для множества целевых классов является процент правильно распознанных наблюдений из обучающей выборки. Для данной задачи ЛДА дал точность классификации 96.8 %, а ИНС – 98.2 % (рис. 5).

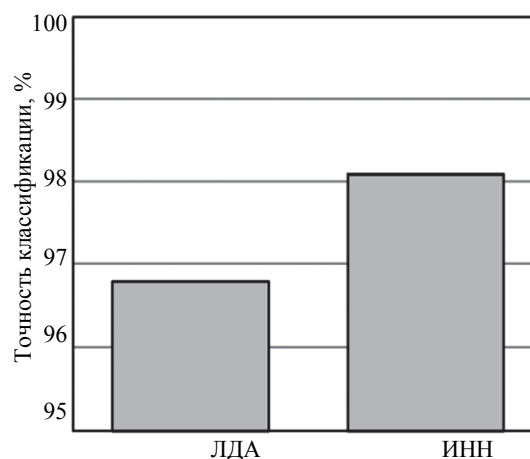


Рис. 5

Как видно из рис. 5, максимально достигнутая точность для ИНС больше, чем для ЛДА. Это связано с тем, что даже простые с точки зрения архитектуры нейронные сети обладают свойством нелинейности и дают результат не хуже линейных классификаторов. Анализ точности классификации для выбора корректной нейронной сети является необходимым, но не достаточным.

На рис. 6 представлены результаты классификации 268 наблюдений по 16 классам для ЛДА и ИНС. Как видно, линейный дискриминантный анализ (рис. 6, а) и классифицирующие нейронные сети (рис. 6, б) дают очень близкий результат классификации. В случае ЛДА точность оказалась меньше из-за того, что в выборке присутствуют два класса, между которыми евклидово расстояние в пространстве признаков минимальное и равно 0.19 (для сравнения, среднее расстояние между классами 1.71, а максимальное 3.65).

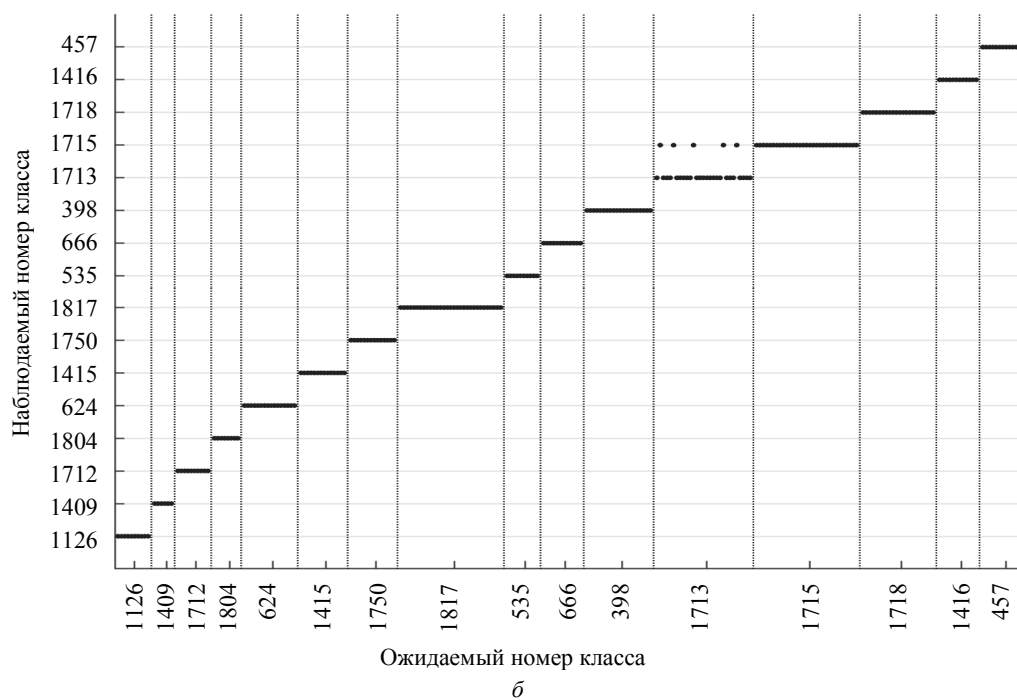
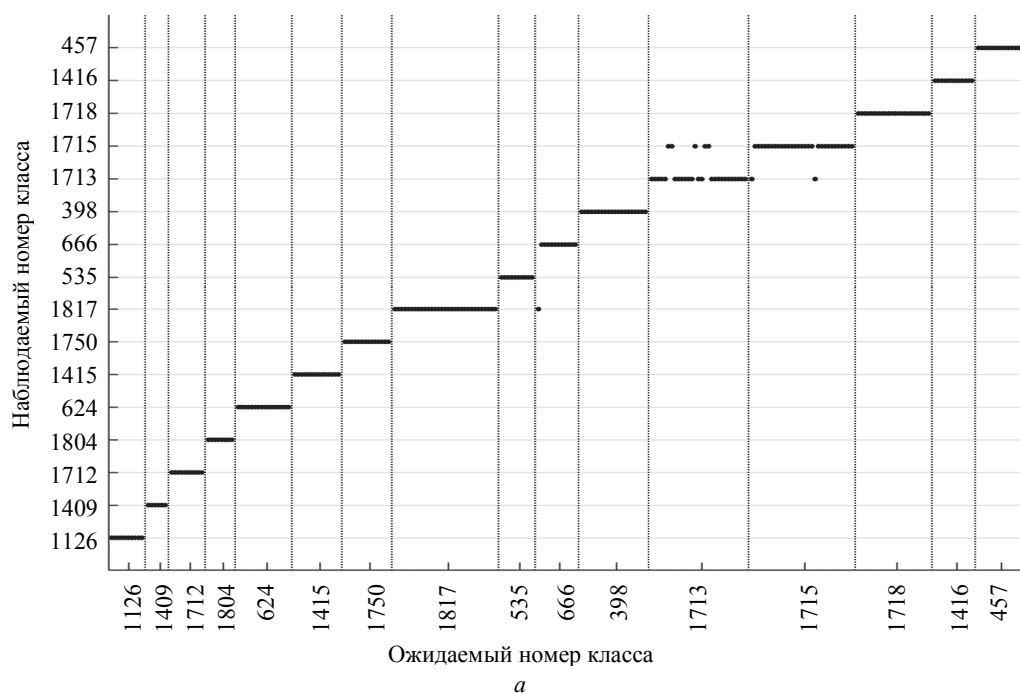


Рис. 6

Подобные результаты свидетельствуют о том, что, с одной стороны, классифицирующие признаки выбраны правильно, но, с другой стороны, на таких сложных данных статистические классификаторы могут давать некоторый сбой. Используя понижение размерности пространства признаков на основе дискриминантного анализа Фишера можно оценить взаиморасположение наблюдений в исходном пространстве признаков. На рис. 7 представлена проекция на трехмерное подпространство, максимизирующее расстояние между клас-

сами, называемое пространством корреляционных функций. Эта проекция на подпространство позволяет исследовать причины ошибок классификации. Например, если происходит ошибка классификации, то по данной диаграмме можно выяснить, происходит ли она с близкими по евклидовому расстоянию классами или нет. В данной работе проекция на трехмерное подпространство использовалась для анализа качества классификации и обобщения нейронной сети, а также для валидации полученной модели нейронной сети.

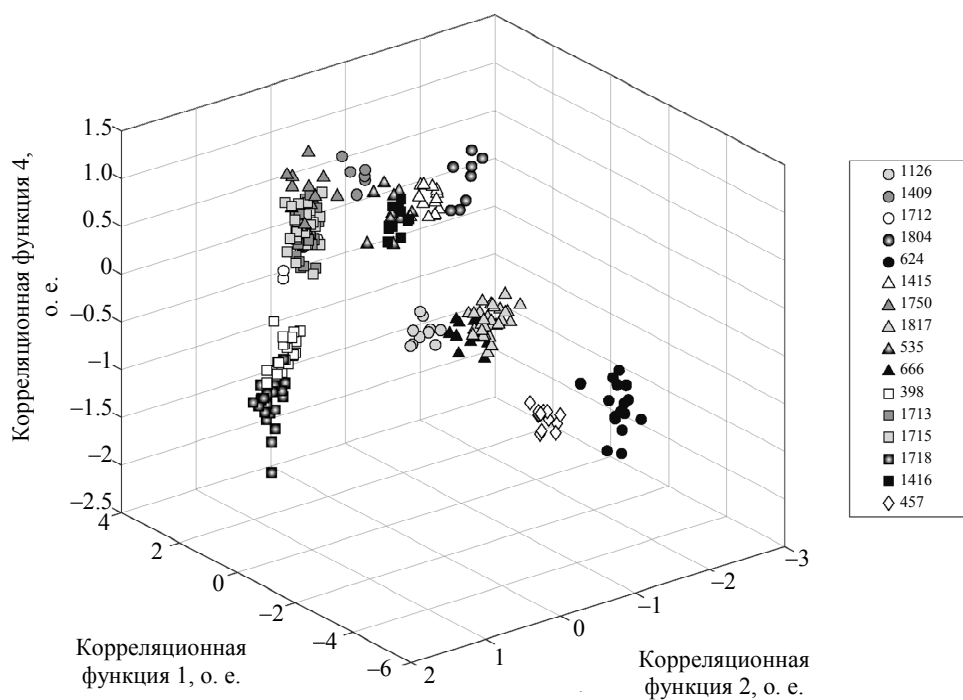


Рис. 7

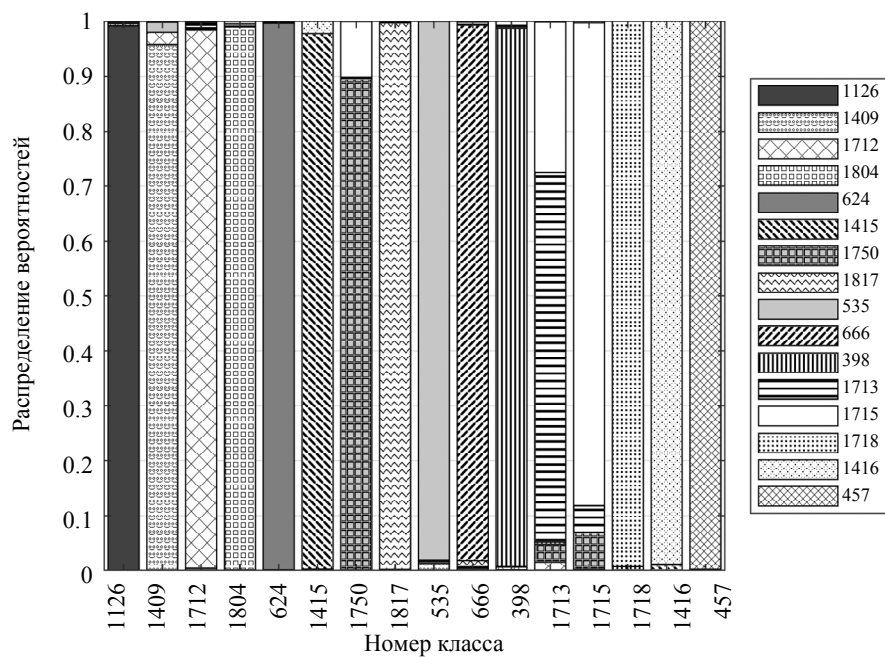


Рис. 8

Ранее говорилось, что используемая нейронная сеть на выходе для каждого наблюдения дает вероятности принадлежности к целевым классам. Класс, имеющий максимальную вероятность, становится классом, к которому принадлежит наблюдение. При таком подходе может возникнуть проблема, когда точность классификации высокая, но максимальная вероятность принадлежности к классу невысокая (менее 60 %). Наличие такой проблемы у сети ведет к тому, что применение конечной модели к наблюдениям, не

принадлежащим к обучающей выборке, может давать ложные результаты. Для выявления данной проблемы необходимо непосредственно анализировать выход нейронной сети. На рис. 8 приведен пример выхода сети, усредненный по наблюдениям внутри каждого класса. Каждый столбец диаграммы соответствует усредненному по классу результату на выходе. Каждый цвет в столбце соответствует одному из 16 целевых классов. Процентное соотношение цветов в столбце показывает распределение вероятности принадлежности к

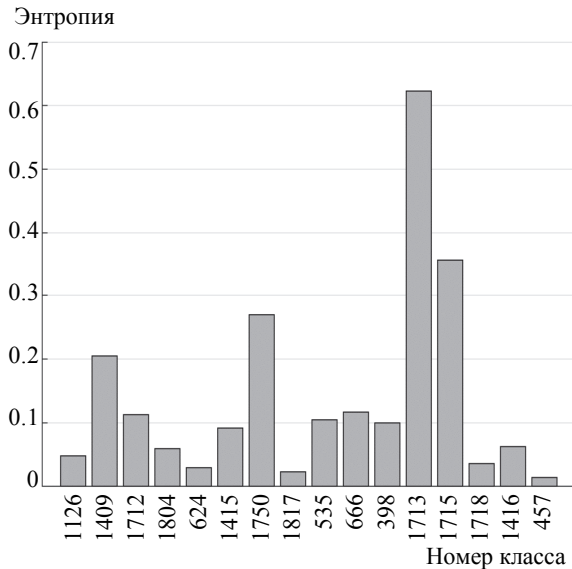


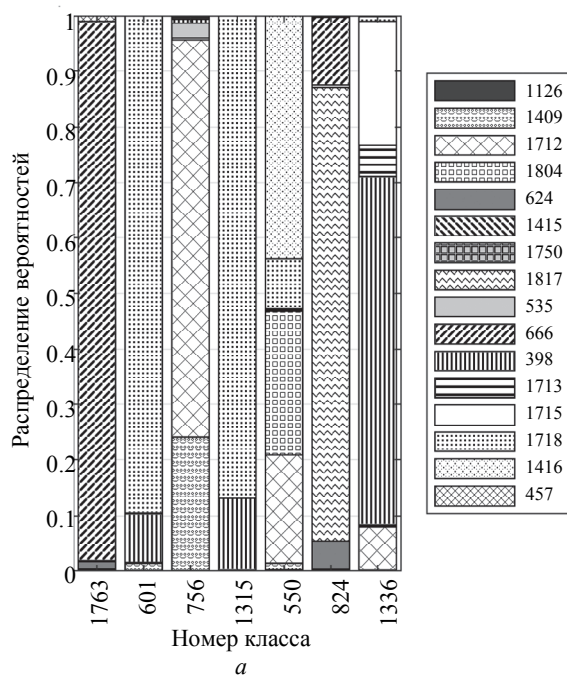
Рис. 9

известным классам. Такая столбчатая диаграмма позволяет оценить правильность классификации по распределениям вероятностей для различных классов, обнаружить корреляцию между классами, а также выявить наличие проблем в классификаторе. Как следует из рис. 8, классы, которые обладают не максимальным значением вероятности на соответствующем выходном нейроне, могут попасть в близкие к ним классы, и в данном случае это может считаться допустимым.

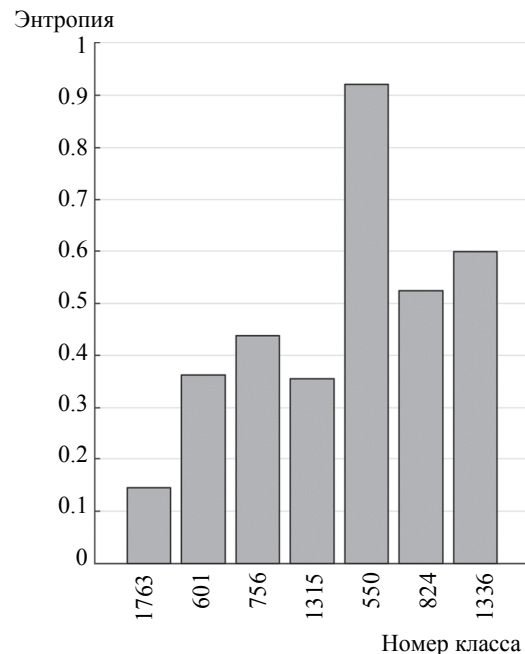
Анализировать выход сети подобным графическим путем затруднительно при большом количестве целевых классов. Поэтому в рассматриваемой работе вводится параметр количественной оценки выхода сети – энтропия, которая показы-

вает степень неравномерности распределения вероятности. В случае нейронных сетей высокая энтропия конкретного класса является показателем того, что для него может возникнуть ошибка классификации. На рис. 9 изображена диаграмма энтропии, усредненной по классам. Диаграмма построена для выхода сети, представленного на рис. 8. Из диаграммы видно, что для двух классов (1750 и 1715), для которых вероятность принадлежности к своим классам была наименьшей, энтропия намного выше, чем у остальных. Следовательно, эти два класса наименее четко определены. Из рис. 6 также видно, что данные классы пересекаются наиболее сильно и не имеют в пространстве признаков четкой границы, которая бы их однозначно разделяла.

Помимо хорошей точности классификации правильно обученная нейронная сеть должна обладать и хорошим качеством обобщения, т. е. правильно определять наблюдения из классов, не присутствовавших в обучающей выборке. Зачастую для анализа качества обобщения применяют тестовую выборку состоящих из наблюдений целевых классов. Такой подход не всегда имеет смысл, особенно, когда выход сети дает высокую вероятность принадлежности к классу. В данной статье предлагается оценивать качество обобщения построенной нейронной сети наблюдениями, которые не принадлежат к целевым классам, но для которых известны максимально близкие классы среди целевых. Информация о близких классах может быть известна априори или быть оце-



а



б

Рис. 10

нена при помощи ЛДА. На рис. 10, а представлен вероятностный выход сети для 7-ми классов, не участвовавших в обучении, и диаграмма энтропии. Как видно из рис. 10, есть группы наблюдений, которые сеть приписывает к нескольким классам с достаточно большой вероятностью. Очевидно, что в общем случае некоторые классы могут располагаться равноудаленно от нескольких целевых классов в пространстве признаков. Высокая энтропия для таких классов показывает неопределенность в классификации данной группы наблюдений (рис 10, б). Естественным недостатком такого анализа обобщения является необходимость априорных знаний о принадлежности и похожести классов. В случае, если априорной информации нет, такой подход может быть полезен для предварительного анализа новых наблюдений.

Результаты оценки качества обобщения также согласуются с результатами проекции на трехмерное подпространство с помощью дискриминанта Фишера (см. рис. 7), новые для нейронной сети наблюдения были отнесены к ближайшим целевым классам. Для тех неизвестных классов,

которые находятся между несколькими известными, вероятностное распределение выхода сети имеет наибольшие вероятности попадания именно в эти ближайшие классы. Заметим, что по соотношению вероятностей нельзя судить об отношении евклидовых расстояний до целевых классов.

Следует отметить, что представленные в работе методики сравнения ЛДА и нейронных сетей применимы для сравнения нейронных сетей с другими алгоритмами машинного обучения и статистическими классификаторами. Эти методики позволяют на раннем этапе построения модели нейронной сети определять ее корректность и качество. Также появляется возможность выбирать наилучшую нейронную сеть из серии построенных, не основываясь только на ошибке классификации и/или ошибке нейронной сети, что позволяет уменьшить риски при внедрении модели в практические приложения. Предложенная методика позволяет анализировать процесс обучения нейронной сети с целью оперативного выявления критичных отклонений в ее работе и возможного алогичного поведения в дальнейшем.

СПИСОК ЛИТЕРАТУРЫ

1. Langford J. Tutorial on practical prediction theory for classification // J. of machine learning research. 2005. Vol. 6, № 3. P. 273–306.
2. Fisher R. A. The Use of Multiple Measurements in Taxonomic Problems // Annals of Eugenics. 1936. Vol. 7. P. 179–188.
3. Goh A. T. C. Back-propagation neural networks for modeling complex systems // Artificial Intelligence in Engineering. 1995. Vol. 9, № 3. P. 143–151.
4. Heermann P. D., Khazenie N. Classification of multispectral remote sensing data using a back-propagation neural network // IEEE Transactions on Geoscience and Remote Sensing. 1992. Vol. 30, № 1. P. 81–88.
5. Huberty C. J. Issues in the use and interpretation of discriminant analysis // Psychological Bulletin. 1984. Vol. 95, № 1. C. 156.
6. Mingoti S. A., Lima J. O. Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms // European Journal of Operational Research. 2006. Vol. 174, № 3. P. 1742–1759.
7. Gabrys B., Bargiela A. General fuzzy min-max neural network for clustering and classification // IEEE transactions on neural networks. 2000. Vol. 11, № 3. P. 769–783.
8. Lawrence S. Face recognition: A convolutional neural-network approach // IEEE transactions on neural networks. 1997. Vol. 8, № 1. P. 98–113.
9. Kochanek D. H. U., Bartels R. H. Interpolating splines with local tension, continuity, and bias control // Acm Siggraph Computer Graphics. ACM. 1984. Vol. 18, № 3. P. 33–41.
10. Comparative performance analysis of hamming, hanning and blackman window / P. Podder, T. Z. Khan, M. H. Khan, M. Rahman // Intern. J. of Computer Appl. 2014. Vol. 96, № 18. P. 1–6.
11. Применение линейного дискриминантного анализа для классификации цианобактерий по спектрам собственной флуоресценции / Т. Р. Жангиров, А. С. Перков, А. А. Лисс, Н. Ю. Григорьева, Л. В. Чистякова // Изв. СПбГЭТУ «ЛЭТИ». 2012. № 5. С. 45–55.
12. Бриллинджер Д. Временные ряды. Обработка данных и теория. М.: Мир, 1980. 536 с.
13. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain // Psychological review. 1958. Vol. 65, № 6. P. 386–407.
14. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002. 344 с.
15. Nesterov Y. Dual extrapolation and its applications to solving variational inequalities and related problems // Mathematical Programming. 2007. Vol. 109, № 2–3. P. 319–344.

T. R. Zhangirov, A. S. Perkov, S. A. Ivanova, A. A. Liss
Saint Petersburg Electrotechnical University «LETI»

N. Yu. Grigoryeva
Saint Petersburg SRCES RAS

L. V. Chistyakova
Center for Culture Collection of Microorganisms SPSU

COMPARISON OF THE EFFICIENCY OF SOLUTION OF CLASSIFICATION TASK BY METHODS OF LINEAR DISCRIMINANT ANALYSIS AND ARTIFICIAL NEURAL NETWORKS

We compare two methods of classification: linear discriminant analysis and artificial feedforward neural networks. The comparison is carried out on the example of the task of classifying cyanobacteria in a series of self-fluorescence spectra of the individual cells. Methods of visualization and analysis of classification results, including new methods based on the entropy of classification results, have been demonstrated. The correlation between the results of linear discriminant analysis and classification of neural networks is determined. The key moments are shown, on the basis of which it is possible to reveal an incorrectly trained model of the neural network, and also a new method for analyzing the quality of the generalization is proposed. The methods of comparing the results of linear discriminant analysis and neural networks presented in the paper are applicable for comparing neural networks with other machine learning algorithms and statistical classifiers. The proposed method allows analyzing the learning process of a neural network in order to quickly identify critical deviations in its operation and possible illogical behavior in the future.

Classification, linear discriminant analysis, neural networks, statistical analysis, gradient descent

УДК 519.852.6

В. Е. Герцман
АО «НИЦ СПб ЭТУ»

А.-В. И. Середа
Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аппроксимация траекторных рядов с использованием кубических В-сплайнов (практический аспект)

Анализируется практическая эффективность использования аппарата кубических В-сплайнов для аппроксимации временных рядов траекторных измерений. Аппроксимирующая функция представляется в виде линейной комбинации базисных сплайнов, обеспечивающей минимум суммы квадратов невязок измерений. Рассматривается возможность аппроксимации рядов с различной степенью сглаженности, согласованная аппроксимация координатных рядов и соответствующих рядов скоростей, интерполяция аппроксимированных данных между узлами исходной сетки. Заполнение пропусков в измерениях осуществляется с помощью линейной модели. Обнаружение интервалов с аномальными измерениями и их корректировка на этих интервалах осуществляется с помощью медианной обработки исходного ряда с последующим сглаживанием с помощью фильтра Хэннинга. Для оценки практической эффективности предлагаемого подхода к аппроксимации траекторных измерений были реализованы алгоритмы расчетов в программе пакете MATLAB и проведена серия вычислительных экспериментов с различными наборами реальных данных.

Временные ряды, кубические В-сплайны, аппроксимация, интерполяция

Под траекторными рядами подразумевается временная последовательность измерений трех составляющих радиуса-вектора и вектора скоро-

сти движущегося высокодинамичного объекта, выполненных бортовым навигационным прибором. Предполагается, что частота измерений су-