

УДК 381.3

С.А. Байрак, В.В. Островский, М.М. Татур

Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Беларусь
tatur@bsuir.by

Тестовая база данных и методика сравнительной оценки классификаторов

Исследована проблема корректной сравнительной оценки классификаторов. В качестве критерия эффективности использованы достоверность (для четких) и точность (для нечетких) классификаторов. Выделены характерные признаки классификации, на основании которых определены 8 типовых задач. Для каждой задачи могут быть вычислены значения функции при случайных либо детерминированных входных данных. Предложена методика постановки экспериментов над классификаторами.

Введение

Классификатор является неотъемлемой частью любой интеллектуальной системы распознавания. В редких случаях он присутствует в «неявном» виде, т.е. не выделяется в самостоятельный модуль, но чаще классификатор выступает в виде отдельного блока. Функции классификатора традиционно состоят в принятии решений и не связаны с подзадачей выделения информативных признаков. Если рассматривать классификатор как черный ящик, то «наблюдаемыми» параметрами являются: число входов, число выходов и способ представления данных (четко/нечетко, фиксированная/плавающая запятая, разрядность и т.п.). Тогда «скрытыми в ящике» характеристиками являются: способ вычисления меры подобия, способ принятия решения. Алгоритм обучения не является непосредственной частью классификатора, но безусловно связан с математикой реализации классификатора и ключевым образом определяет его эффективность. Также обязательной частью исходных данных на проектирование являются обучающие и тестовые выборки.

Задача проектирования классификатора формулируется как выбор (разработка) такого сочетания способа вычисления меры сходства, критерия принятия решения и алгоритма обучения, чтобы точность либо достоверность классификации были максимальными. Иногда дополнительными условиями выступают ограничения на сложность вычислений, время и сходимость обучения, возможность распараллеливания вычислительного процесса при аппаратной поддержке и др.

В работах [1-3] излагалась концепция проектирования систем распознавания, в которой центральное место занимает следующий тезис: *«Классификаторы представляют собой абстрактные автоматы, структура и обучение которых не связаны с конкретной прикладной задачей. Следовательно, разрабатывать и исследовать классификаторы можно и нужно, независимо от конечного применения».*

Для максимального использования накопленного опыта проектирования классификаторов должна быть создана *единая тестовая база и единая методика сравнительной оценки.*

Как правило, в качестве тестовых баз для разработки классификаторов используются частные, интуитивно сформированные, а значит «зашумленные» базы данных. Единой же методики сравнения классификаторов вообще не существует. В настоящей статье приводится проект такой базы данных и методики корректной сравнительной оценки классификаторов, которые могут быть использованы при структурировании и систематизации результатов в области проектирования классификаторов интеллектуальных систем распознавания.

Обоснование базы тестовых данных

Построение базы данных основывается на следующем утверждении: *«В прикладных задачах классификации можно выделить закономерности и сформулировать типовые задачи, структурировать и ранжировать их по “степени интеллектуальности” и показать, как и при каких условиях одна задача вырождается в другую».*

Исходя из этого тезиса, обучающие и тестовые последовательности должны вычисляться, а не формироваться как результат экспертных оценок. Таким образом, можно достоверно знать структуру данных, внутренние логические связи, а соответственно и «сложность». Это даёт мощный механизм, позволяющий ранжировать сами тестовые векторы, соответственно и типовые задачи, а также позволяет определить, какой классификатор сможет лучше справиться с той или иной задачей. Оперируя вычисленными последовательностями, мы сможем получать абсолютные значения достоверности и точности классификации.

В соответствии с изложенными утверждениями, ниже приводятся четыре базовые задачи классификации, различающиеся способами участия информативных признаков в принятии решения.

1. Информативные признаки равноценны.
2. Информационные признаки имеют разные веса в принятии решений.
3. Информационные признаки ограничены (в т.ч. нечетко ограничены) допусками.
4. Информационные признаки логически связаны.

Каждая задача может иметь четкий либо нечеткий вариант принятия решения. Таким образом, число типовых задач составляет 8. Ниже приводится обобщённая модель элементарного классификатора, которая позволяет выразить любую из типовых задач и, следовательно, сгенерировать тестовые векторы соответствующего типа.

$$Y(X) = P\left(\sum_{i=1}^m f_i(x_i) \cdot k_i(x_i)\right) \cdot L\left(\sum_{i=1}^m f_i(x_i) \cdot k_i(x_i)\right),$$

где: m – число информативных признаков,

$f_i(x_i)$ – i -ая функция принадлежности i -го информативного признака,

$k_i(x_i)$ – i -ый весовой коэффициент i -го информативного признака,

$L()$ – функция логического вывода,

$P()$ – пороговая функция принятия решения (чётко/нечётко).

Для наглядности рассматриваются лишь так называемые двухклассовые задачи или задачи идентификации, поскольку вариант многоклассовой задачи может быть сведен к последовательности задач идентификации.

Тестовая база данных представляет собой таблицу, в которой содержатся три варианта по каждой (из 8) типовой задаче с фиксированными параметрами: числом

информативных признаков – 20 и числом тестовых векторов – 100. Информативные признаки нормированы в диапазоне $[0 - 1]$, шаг дискретизации данных унифицирован значением – 0,01.

Методика проведения экспериментов

1. Исходные данные для эксперимента:

1.1. Классификатор (метод классификации), алгоритм обучения классификатора, тестовая база данных.

1.2. Классификатор и алгоритм обучения подстраиваются под количественные параметры тестовой базы данных.

1.3. Эксперименты могут проводиться с ограничением либо без ограничения времени обучения. Например, ограничим время обучения одними сутками.

2. Правила использования тестовой базы: последовательность векторов делится на две части: с 1 по 50 и с 51 по 100.

2.1. Первая часть используется в качестве обучающей последовательности, а тестирование производится на второй части.

2.2. Первая часть используется в качестве обучающей последовательности, а тестирование производится по всей тестовой последовательности.

2.3. Вторая часть используется в качестве обучающей последовательности, а тестирование производится на первой части.

2.4. Вторая часть используется в качестве обучающей последовательности, а тестирование производится по всей тестовой последовательности.

3. В ходе эксперимента оцениваются достоверность классификации (в случаях четкого принятия решений) и точность классификации (в случаях нечеткого принятия решений) по следующим правилам:

3.1. Для каждого из трех вариантов ставится четыре эксперимента согласно 2.1 – 2.4. Таким образом, для каждой задачи ставится $4 \times 3 = 12$ экспериментов.

3.2. Для задач с четким порогом принятия решения рассчитывается достоверность классификации как отношение правильно классифицированных векторов к общему числу векторов, участвовавших в эксперименте. Ошибки первого и второго рода не различаются. Конечный результат по каждой задаче – средняя и минимальная достоверность из 12 экспериментов.

Замечание. При необходимости можно рассчитать достоверность в виде вероятности появления ошибок первого и второго рода. Тестовая база в настоящем виде не содержит подкласс выбросов (нераспознаваемых векторов).

3.3. Для задач с нечетким порогом принятия решения рассчитывается точность классификации как разность между тестовым и полученным результатом классификации по всем векторам. Конечный результат по каждой задаче – средняя и максимальная разность из 12 экспериментов.

Заключение

Результатом проведенных исследований является концепция получения количественных оценок эффективности любых классификаторов. Для этого «насчитывается» база тестовых и обучающих данных под типовые задачи классификации и унифицируется

методика постановки экспериментов с классификатором. Классификатор рассматривается как черный ящик, в котором определены лишь число входов (информативных признаков) и вид пороговой функции.

Логическим развитием данной концепции является:

- 1) проведение состязательного отбора (селекции) как методов классификации, так и алгоритмов обучения в рамках отдельного метода;
- 2) разработка архитектуры универсального (программируемого) классификатора – прототипа нейрокомпьютера;
- 3) удешевление разработки и оптимизация систем распознавания.

Литература

1. Tatur M., Rytickov V., Yarashevich A. The united approach for parallel classifiers design // Proc. of Int Conf. Pattern Recognition and Information Processing. – Minsk. – 2005. – P. 233-235.
2. Татур М.М. Формальные и неформальные аспекты в разработке систем распознавания // Искусственный интеллект. – 2007. – № 3. – С. 333-343.
3. Байрак С.А., Калабухов Е.В., Прытков В.А., Татур М.М. Проблемы корректности сравнительной оценки интеллектуальных систем распознавания // Искусственный интеллект. – 2006. – № 3. – С. 571-574.

С.А. Байрак, В.В. Островский, М.М. Татур

Тестова база даних і методика порівняльної оцінки класифікаторів

Досліджена проблема коректної порівняльної оцінки класифікаторів. Як критерій ефективності використані достовірність (для чітких) і точність (для нечітких) класифікаторів. Виділені характерні ознаки класифікації, на основі яких визначені 8 типових задач. Для кожної задачі можуть бути вираховані значення функції при випадкових або детермінованих вхідних даних. Запропонована методика постановки експериментів над класифікаторами.

S.A. Bayrak, V.V. Ostrovsky, M.M. Tatur

Test Data base and Procedure of Comparative Classifier Rating

The problem of proper comparative classifier rating was researched. Authenticity was taken as efficiency criteria for binary-output classifiers and accuracy – for fuzzy classifiers. Several specific classification signs were highlighted, which resulted in definition of 8 typical tasks. For each of these tasks function value could be calculated both for random and determined input data. Experiment set system for classifiers was proposed.

Статья поступила в редакцию 09.07.2008.