

УДК 004.942

Татур М.М., Крюков А.И.,
Чэнь Цз., Каранкевич В.Г. (Минск, БГУИР)

ОБУЧЕНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПРИНЯТИЯ РЕШЕНИЙ КАК ОБРАТНАЯ ЗАДАЧА ВЫБОРА ПАРАМЕТРОВ МОДЕЛИ

В самом общем виде принятие решений состоит из двух этапов: первый – построение модели F принятия решений $Y=F(X,W)$ в соответствии с прикладной задачей и второй, – собственно, принятие решений по входным данным X . В теории машинного обучения для построения модели F используют различные «унифицированные» математические методы, начиная от классических методов интеллектуального анализа данных (Data Mining) до искусственных нейронных сетей (ANN). При правильном выборе математического метода и посредством настройки параметров W модели F можно достичь приемлемое качество принятия решений, Рис.1. В данной постановке первый этап является *обратной задачей*, а второй этап - *прямой задачей*.

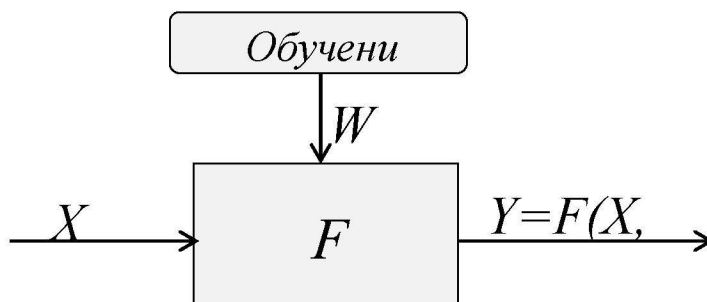


Рисунок 1. – Обобщенная модель принятия решений

Часто, в узкой трактовке, параметры W рассматривают как веса нейронной сети, а процесс вычисления весов – как *обучением нейронной сети*. Гораздо реже параметры W рассматривают как веса вектора информативных признаков X интеллектуального анализа данных. Общепринято считать, что веса признаков (параметры) W , ключевым образом влияют на принимаемые решения. Однако, при более глубоком рассмотрении вопросов влияния весов на принимаемые решения можно заметить существенные различия в механизмах такого влияния. Исследование вопросов влияния весов на принимаемые решения позволяет более обоснованно выбирать модели и интерпретировать результаты интеллектуального анализа данных.

План исследования состоял в следующем. Задана последовательность образов или объектов (Data Set), представленных в виде двумерного вектора количественных информативных признаков. Число и распределение образов могут быть выбраны случайно. Мерность пространства 2 – для простоты расчетов и наглядного анализа результатов исследования. Принятие решения может трактоваться как присвоение метки класса, или кластера в результате алгоритмов классификации, или кластеризации, соответственно. Также «принятие решения» может трактоваться как сравнение двух образов в соответствии с некоторой функцией ранжирования. В ходе проведения исследования решались задачи классификации, кластеризации и ранжирования при различных весовых коэффициентах, различными методами. Образы были индексированы и фиксировались факты изменения меток классов, кластеров либо ранга каждого образа, что свидетельствовало о влиянии изменения веса на принятие решения.

По результатам исследования были сделаны следующие выводы.

В методах классификации, основанных на построении разделяющей или дискриминантной функции, при изменении весов информативных признаков (масштабировании распределения образов) соответствующим образом изменяется положение дискриминантной функции в признаковом пространстве. Поэтому, на результаты классификации (или принятия решений) изменение весов признаков не оказывает.

В методах классификации, основанных на сравнении расстояний между образами, а также в методах кластеризации, на изменение принятия решений оказывает распределение образов, используемая метрика, а также – мерность пространства признаков. Эта зависимость может оказаться достаточно сложной, но очевидно, что наиболее подверженными к изменению решения оказываются образы, находящиеся близко друг к другу в заданной метрике.

При ранжировании, изменение весов информативных признаков наиболее явно влияет на принятие решений, поскольку функции ранжирования, как правило, отражают предметную область и не связаны с формальными метриками оценки расстояний между образами.

Особое место в настоящем исследовании занимают ANN. Безусловно, веса W (обучение ANN) играют ключевую роль в принятии решений, и при этом ANN воспроизводит разделяющую функцию (в задачах классификации). Здесь нет никакого парадокса. В ANN веса нейронов не отражают взаимное (пропорциональное) влияние признаков на принятие решения, т.е. это означает, что настройки W у ANN являются не интерпретируемыми.

Обучение интеллектуальной системы, как решение обратной задачи, связано с большими вычислительными затратами, с наличием различных инвариантов, не гарантирующих оптимального решения. Принимая во внимание, что для каждой модели F будут характерны свои, уникальные сложности в вычислении или подборе параметров W , то это означает актуальность продолжения исследований в данной области для молодых (и не только) специалистов в области DataScience.