АНАЛИЗ ДАННЫХ

Автор: словарь
31.08.2007 15:50 -

АНАЛИЗ ДАННЫХ - область математики и информатики, занимающаяся построением и исследованием наиболее общих математических методов и вычислительных алгоритмов извлечения знаний из экспериментальных (в широком смысле) данных.

В социологии А.Д. применяется в основном к анализу данных опросов. Первоначально А.Д. практически сводился к прикладным разделам математической статистики, поскольку объектом анализа являлись в основном случайные выборки из тех или иных совокупностей, называемых генеральными, что и составляет предмет математической статистики.

Прикладная математическая статистика и в настоящее время представляет собой наиболее обширный раздел А.Д.

Вместе с тем необходимость исследования больших массивов данных, не являющихся случайными выборками, - например, содержимого баз данных и данных в сети Интернет - привели к созданию других подходов, из которых в первую очередь следует отметить так называемый Data Mining (DM), что может быть переведено как "раскапывание данных" (в русскоязычной литературе термин обычно используется без перевода). DM в настоящее время несколько эклектичен по набору используемых методов и представляет собой скорее некоторую идеологию. Разницу между математической статистикой и DM можно проиллюстрировать следующим образом.

Если типичной задачей математической сатистики является задача "найти зависимость между заданными переменными", то для DM характерны задачи типа "найти переменные, между которыми существует достаточно хорошо выраженная зависимость". DM находится в стадии интенсивного развития как в плане расширения арсенала используемых методов, так и углубления концепции, что в перспективе может привести к цельной системе методов, различающихся по шкале "направленный анализ - поисковый анализ". Однако и на нынешнем этапе развития DM может успешно применяться в социологии. Имеются и принципиально иные разделы А.Д., пока еще в социологии практически не используемые, - например, математическая морфология, связывающая методы количественного анализа и визуального представления данных, или теория интерпретации эксперимента, которая сближает традиционный А.Д. и моделирование математическое (см.).

Н.Н. Леонов