

ПРЕДИСЛОВИЕ

Классификация — это мощное средство структурирования сложных систем, возникающих в науке, технике, обществе и любых других видах человеческой деятельности. Математическая классификация — это дисциплина, связанная с разработкой математических методов построения и анализа классификаций. Как научная дисциплина математическая классификация находится на стыке математической и прикладной статистики, информатики и теории искусственного интеллекта, дискретных и непрерывных задач оптимизации.

За последние тридцать лет исследования в области математической классификации проводились в самых различных теоретических и прикладных дисциплинах, нередко с использованием сильно различающейся терминологии. Так, в технических исследованиях говорят о задачах распознавания образов, задачах технической диагностики, нейронных сетях, анализе текстов и изображений, декомпозиции сложных систем и пр., в задачах моделирования психических процессов говорят о машинном обучении, кластерном анализе, латентно-структурном анализе, многомерном шкалировании и пр., в биологии занимаются нумерической таксономией, кладистикой (эволюционной таксономией), построением филогенетических древ, и пр.

Создаются национальные и международные научные общества, призванные координировать исследования в данной области; эти общества имеются практически по каждому из вышеуказанных направлений, а также по многим не названным, хотя и близким дисциплинам (например, по теории и практике размытых множеств, по библиотечным классификациям, и т. д.).

Начало координации подобных работ в России и странах СНГ относится к середине 70-х годов и связано с деятельностью возглавлявшегося профессором С. А. Айвазяном Комитета по прикладным методам математики и вычислительной технике Всесоюзного совета научно-технических обществ. Вопросам математической классификации было уделено существенное внимание на заседаниях организованных этим Комитетом и Центральным экономико-математическим институтом Российской Академии наук пяти школ-семинаров «Программно-алгоритмическое обеспечение прикладного многомерного статистического анализа» (пос. Цахкадзор, Армения: 1979, 1983, 1987, 1991 и

1995 гг.), а также пяти конференций «Применение многомерного статистического анализа в экономике и оценке качества продукции» (пос. Кяярику, Эстония: 1977, 1981, 1985 и 1989 гг.; Москва — 1993 г.). В середине 80-х годов в рамках упомянутого Комитета создается специальная «Комиссия по классификации». По ее инициативе и при организационной поддержке Пушкинского биологического центра в декабре 1990 г. в Пушкино-на-Оке проводится Всесоюзная конференция по теории и практике классификации (с приглашением зарубежных специалистов). По инициативе и под редакцией ведущих отечественных специалистов переводятся на русский язык монографии и сборники зарубежных авторов, посвященные математическим методам классификации, среди которых отметим изданную под редакцией С. А. Айвазяна и В. М. Бухштабера коллективную монографию французских ученых «Optimisation en Classification Automatique» (в русском варианте — «Методы анализа данных: подход, основанный на методе динамических сгущений»; изд-во «Финансы и статистика», Москва, 1985 г.). Кстати, в предисловии к этой монографии С. А. Айвазян и В. М. Бухштабер излагают подход к построению общей теории автоматической классификации, доработанный вариант которой был затем доложен на I Всемирном конгрессе общества Бернулли по теории вероятностей и математической статистике (Ташкент, 1986 г.).

Публикацию данного выпуска мы рассматриваем как продолжение этой активности.

Представленные статьи достаточно репрезентативно показывают основные направления российских работ в области математической классификации.

Обзорная статья Абусева, Лумельского посвящена построению асимптотически оптимальных решающих правил поточечной и групповой статистической классификации в случае многомерного нормального распределения и распределения Уишарта. В статье Бородкина рассматривается метод размытого распознавания, ориентированный на анализ экономического развития регионов России сто лет назад, на пороге ее вступления в капитализм — тематика неувядающей актуальности. В статье Грушо изучается сходная задача, но для ситуации, когда данные представлены функцией близости между объектами. В статье Коваленко методика поиска кластеров «высокой плотности» распространяется и обосновывается для случая переменного порога существенности. В статье Кузнецова и Финна рассматривается технология классификации и прогнозирования, основанная на весьма общих математических формализациях концепций «сходство» и «уровень обобщения». В статье Миркина анализируется связь между двумя подходами к математической классификации, ассоциируемыми, прежде всего, с так называемыми нумерической таксономией, с одной стороны, и концептуальным кластерингом, с другой. Статья Моттля, Мучника, Ивановой и Блинова может

рассматриваться как далеко идущее обобщение техники анализа смесей распределений, предназначенное для обработки изображений и другой подобной информации (в данном случае — интерпретации данных сейсморазведки).

Редакторы-составители надеются продолжить работу по публикации трудов специалистов в области математической классификации, работающих на территории бывших советских республик. За справками обращаться к А. П. Коваленко (E-mail: kov@novelta.msk.ru) или в редакции научного изд-ва «ТВП».

Б. Г. Миркин, А. П. Коваленко