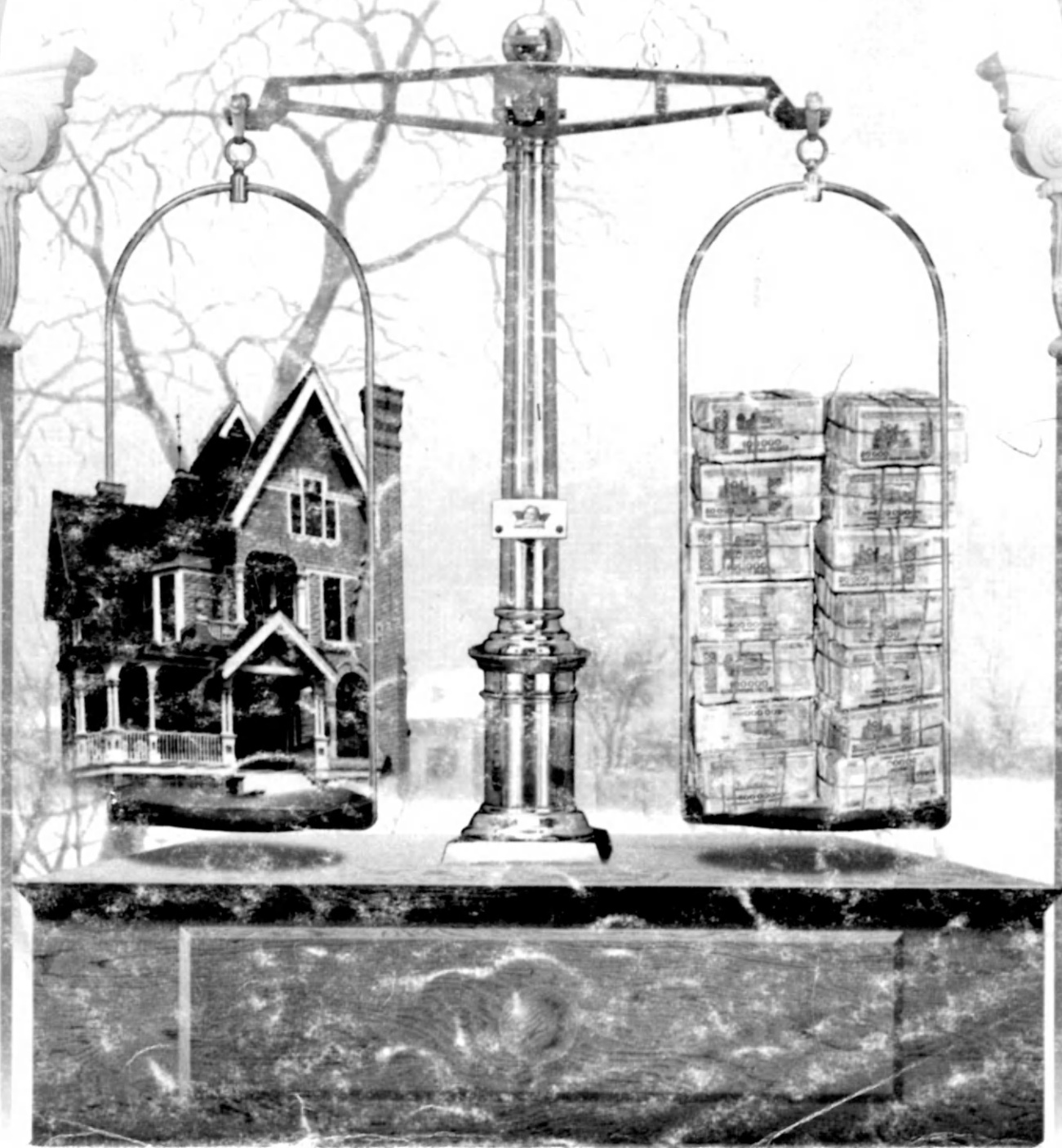


ЖКР-1
227

1999. №4

Банкаўскі • 4/99
ВЕСНІК *Апрэль*

**Залог
как форма
обеспечения кредита**





Александр КУЛЕШОВ,

доцент кафедры уравнений
математической физики
Белгосуниверситета,
ведущий научный сотрудник
НИЛ информационных
систем ЦИТ БГУ



Юрий ПОЗНЯК,

заведующий НИЛ
информационных систем
ЦИТ БГУ



Сергей ЗЕМСКОВ,

младший научный сотрудник
НИЛ информационных
систем ЦИТ БГУ

Современные технологии интеллектуального анализа экономических процессов на основе компьютерной математики

Системы интеллектуального анализа данных (ИАД) реализуют новую форму анализа данных, основанную на интеллектуальном подходе. Их назначение — автоматически выявлять неочевидные для пользователя правила и закономерности, которые помогают оптимизировать процесс принятия управленческих решений в самых ответственных областях деятельности.

Рынок систем ИАД сегодня активно развивается. По данным исследовательской компании Meta Group, в 1997 году его объем достиг 300 миллионов долларов, а к 2000 году он вырастет до 800 миллиардов. В качестве основных причин, способствующих распространению этой технологии, можно назвать следующие:

- осознание того, что в больших по объемам базах данных содержатся не “горы” бесполезной информации, а скрытые ценные знания, характеризующие экономические процессы и способные в конечном счете повысить эффективность управления;
- развитие технологии информационных хранилищ (DATA Warehousing), поскольку с точки зрения аналитики гораздо удобнее иметь дело с единым информационным пространством, собрав требуемые для анализа данные в центральной БД, очистив их от ошибок, приведя к единым форматам и представив в удобном для пользователя-аналитика виде;
- снижение стоимости устройств хранения информации;
- уменьшение стоимости вычислительной техники;
- обострение конкурентной борьбы за клиента;
- переход от массового обслуживания к сегментному и индивидуальному и так далее.

В процессе пользования ИАД могут выполняться различные операции, которые, в свою очередь, реализуются при помощи разнообразных алгоритмов. В основе большинства из них лежит аппарат современной математической статистики. Отдельный класс составляют методы нечеткой логики и основанные на этих методах экспертные системы.

Если необходимость использования ИАД сейчас уже не вызывает сомнений, то на первый план выходит главная проблема: каким образом производить внедрение ИАД, в частности, в сферу банковской деятельности для наиболее полного использования всей мощи этой новой технологии.

Самый простой способ — закупка разрозненных пакетов и эксплуатация их по технической документации — оказывается малоэффективным по нескольким причинам.

Первая причина состоит в том, что заформализовать технологию интеллектуального анализа данных практически невозможно, а сам анализ данных не только наука, но и искусство. Иногда приходится браться за решение задачи вслепую, не “обложившись” теоретическими доказательствами и обоснованиями, а изыскивая ответы на вопросы из компьютерных экспериментов. Так как наши познания всегда ограничены, то, как правило, возникают модели с неполной информацией или даже не имеющие полностью формализованной математической постановки. Кроме того, любой математический пакет строго формализован выбором конечного числа фиксированных вариантов для любой возможной ситуации, возникающей на практике, и уже только поэтому выводимые правила в большинстве случаев не могут привести к правильному решению.

Остановимся на целесообразности применения статистических пакетов более подробно. В современных статистических пакетах наиболее широко представлены следующие методы прогнозирования временных рядов:

1. Методы прогнозирования с помощью средних арифметических.
2. Метод простых средних.
3. Метод двойных движущихся средних.
4. Методы прогнозирования с помощью экспоненциального усреднения.
5. Адаптивный метод Тригга и Лича.
6. Линейный метод Брауна.
7. Линейный метод Холта.
8. Квадратический метод Брауна.
9. Метод Уинтерса для экстраполяции динамических рядов с сезонными колебаниями.
10. Методы авторегрессии и скользящего среднего.

Временной ряд — совокупность последовательных измерений значений переменной (процесса), произведенных через определенные (чаще всего равные) интервалы значений параметров (чаще всего времени). Для анализа получаемых при таких измерениях статистических данных предназначена область статистики, называемая анализом временных рядов.

Анализ временных рядов используется, в частности, для решения следующих задач:

- для построения математической модели процесса, представленного временным рядом, например, модель динамики курса белорусского рубля;
- для исследования структуры временного ряда, например, для выявления изменения среднего уровня значений (тренда) и обнаружения периодических колебаний (например, для исследований сезонных колебаний процентных ставок рынка однодневных межбанковских кредитов);
- для прогнозирования будущего развития процесса, представленного временным рядом, например, для прогнозирования объемов наличных денег в обращении;
- для исследования взаимодействий между различными временными рядами, к примеру, количественного описания зависимо-

сти динамики денежной массы от валового внутреннего продукта.

Как правило, исходной информацией является набор чисел X_1, X_2, \dots, X_N , который представляет собой N наблюдений одной и той же величины. Этот набор и называется временным рядом. Прогнозирование — это попытка предсказать, какое значение интересующей нас величины будет получено при $N+1$ наблюдении, $N+2$ наблюдении и, вообще, при $N+m$ наблюдении.

Основное отличие методов 1—9 и методов 10 — это применимость их в зависимости от длины временного ряда. На коротких временных рядах методы 1—9 применимы, а методы 10 — нет. Методы 1—9 основаны на “наивном” подходе, связанном с вычислением средних арифметических и экспоненциальном сглаживании, а методы 10 основаны на достаточно тонких теоремах из теории вероятности, теории функций комплексного переменного, функционального анализа. Поэтому возникают достаточно четко обозначенные границы применимости методов 1—9 и методов 10. Эти границы колеблются в пределах 30—50 отсчетов временного ряда. Если быть более точным, то из статистической практики известно, что при длине временного ряда менее 30 отсчетов применять более или менее обоснованные теоретически статистические процедуры нельзя, поскольку их обоснование носит асимптотический характер и базируется на предположении о стационарности (незначительности отклонений значений ряда от средней величины).

В то же время для экономического анализа в странах с переходной экономикой представляют интерес в основном “короткие” временные ряды. Фактически это означает, что требуется спрогнозировать ситуацию по непродолжительному (недавно возникшему) процессу. Это и есть основная причина рискованности применения в этих целях инструментов, использующихся в статистических пакетах. Что касается методов 1-9, то, хотя они и применимы на коротких последовательностях данных, получаемые с их помощью результаты можно считать более чем сомнительными. В приложении, помещенном в конце статьи, в качестве

примера альтернативного подхода к прогнозированию коротких временных последовательностей авторы обсуждают вариант прогноза курса доллара США по отношению к белорусскому рублю, основанного на использовании дискретных ортогональных полиномов.

Вторая причина малозффективности эксплуатации стандартных статистических пакетов состоит в том, что использование наукоемких технологий с нулевого уровня невозможно. Для того чтобы подняться на следующую ступень познания, нужно пройти все предыдущие, сохраняя накопленный опыт. Это возможно реализовать только создавая научные коллективы, время от времени пополняемые новыми кадрами, которым передается весь накопленный опыт и которые одновременно проходят всю технологическую цепочку обучения. При этом использование готовых продуктов возможно только как вспомогательный инструмент для пополнения базы знаний о предметной области и отбора рабочих данных, которые будут подвергнуты анализу. Как следует из проведенного выше анализа прогнозирования временных рядов, актуальной на сегодняшний момент является разработка обоснованных процедур прогноза коротких временных рядов. Это одна из задач, которая могла бы стать предметом разработки в научном коллективе.

Инструментарием для этого могут служить пакеты статистической обработки данных разрозненных производителей. Однако при этом возникают чисто технические трудности, потому что исходные тексты коммерческих программ, как правило, недоступны. Речь идет не о научном наполнении, которое все равно нужно адаптировать к возникающим задачам, а об отсутствии объединяющей платформы.

Таким образом, третья существенная причина нецелесообразности использования готовых программных продуктов — отсутствие единой платформы, на которой можно было бы адаптировать, эксплуатировать и развивать готовые пакеты ИАД.

Проанализировав эти проблемы, авторы пришли к выводу, что наиболее оптимальной является следующая технология интеллектуального анализа данных:

1. Выбор интеллектуальной платформы с возможностями объектно-ориентированного программирования и основательной математической базой знаний — своего рода математикой внутри компьютера.

Такие пакеты — это системы компьютерной математики Matlab, Mathematica и другие. Однако пакет Mathematica выделяется тем, что регулярно снабжается надстройками необходимой для ИАД ориентации. Приведем необходимый перечень. Это, естественно, сам пакет Mathematica, являющийся платформой, со следующим минимальным насыщением:

- Experimental Data Analyst
- Time Series
- Fuzzy Logic
- Technical Trader
- Finance

Этот мощный базовый набор позволяет достаточно полно производить интеллектуальный анализ данных при условии наличия моделирующих алгоритмов. Обучение моделированию будет, по нашему мнению, составлять основу учебного процесса.

2. Обучение специалистов работе на этой платформе с соблюдением следующих условий:

- организация и финансирование подготовки специалистов должны быть ориентированы на моделирование сложных экономических процессов;
- учебный цикл должен быть органично связан с тем базовым инструментарием, который они будут использовать в практической деятельности;

● весь процесс обучения должен строиться на принципе теория — модель — реализация модели на базовом инструментарии.

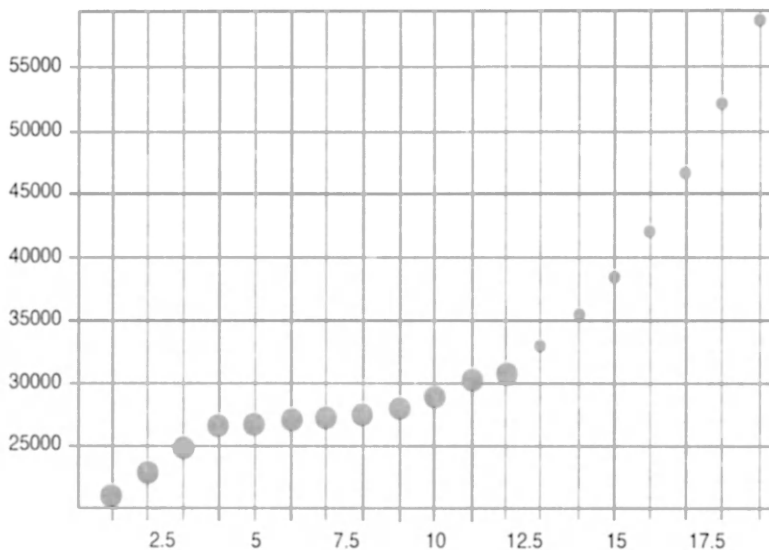
3. Поиск новых знаний, поступающих через периодические научные издания, Internet, научные монографии с целью их обработки и внедрения как в учебный процесс, так и в практическую деятельность будущих специалистов.

4. Использование практических задач организаций-заказчиков в качестве основы тем курсовых и дипломных проектов.

В качестве примера анализа данных с помощью компьютерной математики рассмотрим пример прогнозирования курса доллара США по отношению к белорусскому рублю на семь месяцев (январь — июль) 1998 года.

Прогноз осуществлен в Trial version системы компьютерной алгебры Mathematica 3.0.1 (лицензия № L1080-2714). Исходными данными послужили помесечные значения курса за 1997 год ("Экономические тенденции в Беларуси", февраль 1998 г.).

Сравнительный график фактических значений курса за 1997 год (крупные точки) и расчетных значений для 1997 года и января — июля 1998 года (мелкие точки)



Вертикальная ось — значения курса белорусского рубля по отношению к доллару США. Значения горизонтальной оси от 1 до 19 соответствуют девятнадцати месяцам, с января 1997 года по июль 1998 года.

УВАЖАЕМЫЕ ГОСПОДА!



Банковский вестник, 4' 1999

Фирма "Ариадна"
предлагает

БАНКОВСКОЕ ОБОРУДОВАНИЕ И ГАРАНТИРУЕТ ЕГО ОБСЛУЖИВАНИЕ

- Машина для счета купюр фирмы NGZ.
- Детекторы валют BellCop.
- Вакуумная машина для упаковки банкнот в полиэтиленовые пакеты Multivak.
- Уничтожитель бумаг Crocodile.
- Счетчик монет NGZ.
- Машина для обандероливания Indupro.

Предоставляем гарантию 12 месяцев, продажа осуществляется за безналичные рубли по курсу МВБ. Осуществляем послегарантийное обслуживание, поставку расходных материалов и запчастей по отдельному контракту. Заключаем договора на ежемесячное обслуживание банковского оборудования.

ОБРАЩАТЬСЯ ПО АДРЕСУ:

220030, г. Минск, ул. К. Маркса, д. 20,
телефакс: (0172) 2104076, 2292844

