

УДК 68.13

**Р. Н. Кветный, д. т. н., проф., В. Ю. Коцюбинский, к. т. н., А. Н. Козачко, к. т. н.,  
Л. Н. Кислица, Н. В. Казиминова**

## **РАЗРАБОТКА СПОСОБНОЙ К САМООБУЧЕНИЮ СИСТЕМЫ ПРИНЯТИЯ РЕШЕНИЙ**

*Разработано структурную схему и модель принятия решения для системы, способной самообучаться и вырабатывать пользователю советы для работы на финансовых рынках*

**Ключевые слова:** самообучение, система принятия решений, дерево решений, нечеткая логика

### **Актуальность**

Развитие технологии и повышения производительности труда во всех отраслях относятся к важнейшим заданиям технического прогресса нашего общества. Решение этих заданий возможно лишь при широком внедрении систем автоматического регулирования и управления как отдельными объектами, так и производством в целом. Системы автоматического управления необходимы для управления разными объектами, в частности системами с переменными параметрами. Существует большое разнообразие систем адаптивного управления, однако целью этого исследования являются самостоятельно обучающиеся системы для работы на финансовых рынках с помощью Data Mining [1].

Сегодня все больше возрастает необходимость в системах, которые способны не только выполнять однажды запрограммированную последовательность действий над предварительно определенными данными, но и способные сами анализировать информацию, которая поступает, находить в ней закономерности, делать прогноз и т.д. В наше время накоплен большой опыт создания автоматических систем управления в разных отраслях народного хозяйства. Этот опыт позволяет сделать вывод о том, что повышение эффективности управления заключается в увеличении уровня интеллектуализации этих систем, переходе к так называемым "умным" производственным системам, ориентированным на знание. Сфера применения систем искусственного интеллекта, которые существуют на сегодняшний день, охватывает медицинскую диагностику, интерпретацию геологических данных, научные исследования в химии, биологии, военном деле и другие отрасли.

Способная к самообучению система автоматического управления (самонастраивающаяся система), в которой приспособление к условиям, которые случайно изменяются, обеспечивается автоматическим изменением параметров настройки или путем автоматического поиска оптимальной настройки. В любой другой автоматической системе управления есть параметры, которые влияют на стойкость и качества процессов управления и могут быть изменены при регуляции (настройке) системы. Если эти параметры остаются неизменными, а условия функционирования (характеристики управляемого объекта, возбуждающие действия) существенно изменяются, то процесс управления может ухудшиться или даже стать неустойчивым. Ручная настройка системы часто оказывается сложной, а иногда и невозможной. Использование в таких случаях самообучающейся системы технически и экономически целесообразно и даже может оказаться единственным способом надежного управления [1].

В тех случаях, когда самонастройка применяется в системах управления в следствие недостоверности знания свойств объекта, система самооптимизации напоминает процесс самообучения. Система при этом путем автоматического поиска словно бы сама познает

неизвестные свойства управляемого объекта и учится управлять этим объектом наилучшим образом. Применение таких систем выгодно в тех случаях, когда внешнее действие может быть измерено с целью анализа его свойств и когда изменение его формы является решающим для качеств работы системы.

Особенность системы заключается в том, что она со временем будет настраиваться, обучаться и накапливать опыт пользователя. Актуальность самообучающейся системы именно в том, что человек не вмешивается в настройку и ее обучение. В системе, которая будет разрабатываться в данной работе, экспертная оценка будет индивидуальной для каждого инвестора, его набора факторов и полученных трейдов.

### **Цель исследования**

Целью исследования является разработка автоматизированной экспертной системы для выработки советов пользователю, которая имеет способность к самообучению и на основе постоянно пополняемых данных способна накапливать опыт в принятии решений. Разработка такой системы даст возможность повысить эффективность работы на фондовых рынках и прибыль инвестора.

### **Постановка задачи**

В работе поставлено задание по разработке автоматизированной системы, способной к самообучению и выработке советов инвестору в виде экспертных решений, построения модели принятия решений и методики ее работы.

Решение этой задачи целесообразно разбить на несколько этапов:

1. Анализ и исследование начальной выборки данных.
2. Разработка модели принятия экспертного решения.
3. Выработка совета в зависимости от полученных результатов классификации.

Работоспособность автоматизированной системы в значительной мере зависит от модели принятия решения. Модель принятия решения будем искать в таком виде

$$Y = f(X_1, X_2, X_3, X_4), \quad (1)$$

где  $Y$  – результат принятия решения;  $X_1$  – положение цены относительно скользящего среднего;  $X_2$  – следование ценовой модели;  $X_3$  – изменение процентной ставки;  $X_4$  – субъективное состояние трейдера.

В табл. 1 приведено описание характеристик факторов  $X_1$ -  $X_4$ . Из этой таблицы видно, что факторы  $X_1$  и  $X_2$  представляют собой лингвистические переменные, значения которых вносят в начальные данные неопределенность.

Таблица 1

## Характеристики факторов модели

| Фактор | Тип             | Значения                                                                                                                                                                                         |
|--------|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $X_1$  | Лингвистический | «Н» – выше скользящего среднего в пределах 1%<br>«МН» – выше скользящего среднего более чем 1%<br>«L» – ниже скользящего среднего в пределах 1%<br>«ML» – ниже скользящего среднего более чем 1% |
| $X_2$  | Лингвистический | «SAE» – симметричная модель с растущими торговыми объемами<br>«NAE» – несимметричная модель с растущими торговыми объемами<br>«N» – несимметричная модель                                        |
| $X_3$  | Количественный  | «0» – ставка неизменная<br>«-0.25» – снижение ставки на 0.25%<br>«-0.5» – снижение ставки на 0.5%<br>«0.25» – повышение ставки на 0.25%<br>«0.5» – повышение ставки на 0.5%                      |
| $X_4$  | Дискретный      | «1» – покупать<br>«-1» – не покупать<br>«0» – вне рынка                                                                                                                                          |

Будем считать, что результат принятия решения принимает такие значения

$$Y = \begin{cases} class1 - \text{меньше 1\% прибыли} \\ class2 - \text{больше 1\% прибыли} \\ class3 - \text{меньше 1\% убытка} \\ class4 - \text{больше 1\% убытка} \end{cases}$$

Таким образом построение модели принятия решения (1) сводится к решению задачи классификации.

Для решения задачи классификации с неопределенными начальными данными на практике наибольшее распространение получил метод дерева решений, который позволяет определить набор правил классификации типа «Если - То». Эти правила и будут советом системы пользователю. Рассмотрим основные положения метода дерева решений.

### Основные положения метода дерева решений

Метод дерева решений – это один из методов автоматического анализа огромных массивов данных [2]. Первые идеи создания деревьев решений начинаются с работ Ховленда и Ханта конца 50-х годов XX века. Однако основополагающей работой, что дала импульс для развития этого направления, стала книга Ханта, Мерина и Стоуна «Experiments in Induction», которая была опубликована в 1966р.

Область использования метода дерева решений можно объединить в три класса:

- описание данных: деревья решений позволяют хранить информацию о выборке данных в компактной и удобной для обработки форме, что содержит в себе точные описания объектов;
- классификация: деревья решений прекрасно справляются с задачами классификации, то есть соотношение объектов к одному из описанных классов;

- регрессия: если переменная имеет недостоверные значения, то деревья решений позволяют определить зависимость этой целевой переменной от независимых (входных) переменных.

На рис. 1 приведен пример дерева решений. Это дерево представляет собой совокупность узлов и листьев, которые на рисунке обозначены овалами и прямоугольниками. В узлах записанные условия (тесты), а в листьях – метка класса.

Для принятия решения с помощью дерева решений необходимо выполнить такие шаги:

1) оценить состояние рынка вектором факторов  $X=(X_1, X_2, X_3, X_4)$  (Этот шаг выполняется пользователем системы);

2) определить класс роста прибыли путем движения вектора  $X$  по дереву решений с верхних уровней до нижних (Этот шаг выполняется системой).

На сегодняшний день существует значительное количество алгоритмов, которые реализовывают деревья решений CART, C4.5, NewId, IRule, CHAID CN2 и т.д. В данной работе для разработки модели принятия экспертного решения будет использоваться алгоритм C4.5 [3].

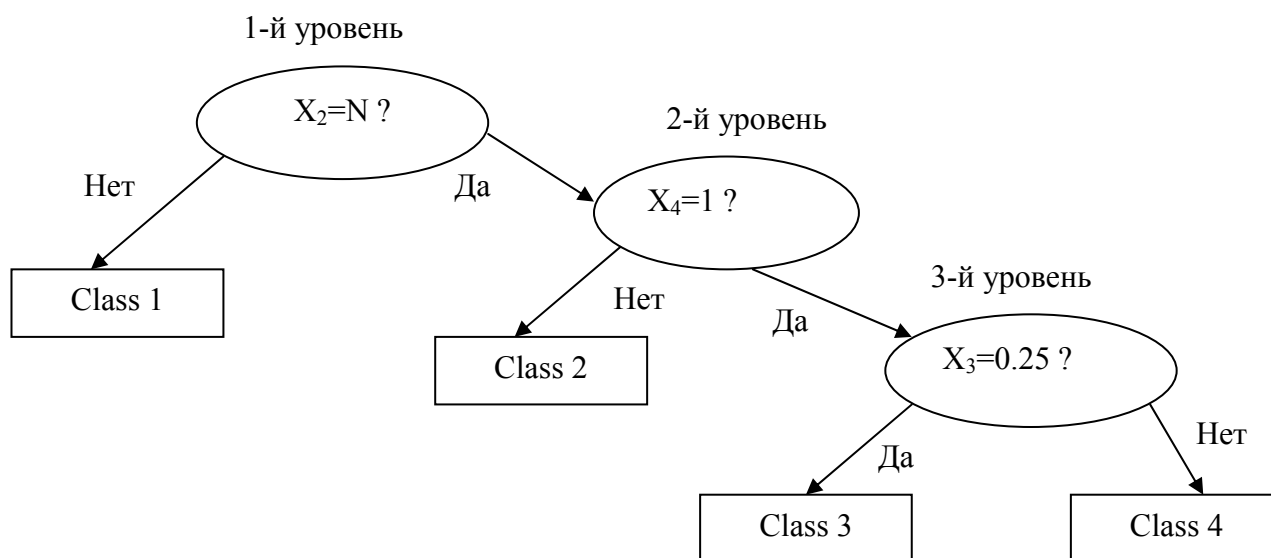


Рис. 1. Пример дерева решений

### Основные положения алгоритма C4.5

Синтез дерева решений сводится к такой задаче: известная выборка инвестора

$$T = \{X^i, Y^i\}, i = \overline{1, n}, \quad (2)$$

где  $X^i = \{X_1, X_2, X_3, X_4\}$  – вектор состояния рынка, сформированный инвестором в виде  $i$  -ой оценки;  $Y^i \in \{class_1, class_2, class_3, class_4\}$  – реальный результат принятия решения для  $i$  -ой оценки инвестора;  $n$  – количество оценок, инвесторов рынка (размер выборки).

Необходимо с помощью сформированной начальной выборки построить функцию

$$X_1 \times X_2 \times X_3 \times X_4 \rightarrow Y, \quad (3)$$

которая превращает пространство факторов в прогнозирование класса роста прибыли.

Для построения дерева на каждом внутреннем узле необходимо найти такое условие (проверку), какое бы разбивало выборку, ассоциируемую с этим узлом на подмножества. В качестве такой проверки должен быть выбран один из атрибутов (факторов). Общее правило для выбора атрибута можно сформулировать таким образом: выбранный атрибут должен разбить выборку так, чтобы получаемые подмножества состояли из объектов, которые принадлежат к одному классу, или были максимально приближены к этому, то есть количество объектов из других классов в каждом из этих подмножеств было как можно меньше. В алгоритме C4.5 предлагается следующий критерий отбора [3]:

$$Gain(X) = Info(T) - Info_X(T), \quad (4)$$

где  $Info(T)$  – энтропия общей выборки;

$Info_X(T)$  – энтропия фактора X, которая вычисляется по формуле:

$$Info_X(T) = \sum_{i=1}^n \frac{|T_i|}{|T|} * Info(T_i). \quad (5)$$

### Структура автоматизированной системы

На рис. 2 приведена структурная схема системы, способной к самообучению и выработке советов пользователю. Эта система состоит из трех блоков:

- база оценок состояния рынка;
- классификация оценки методом дерева решений;
- автоматизированная подсистема принятия решений.

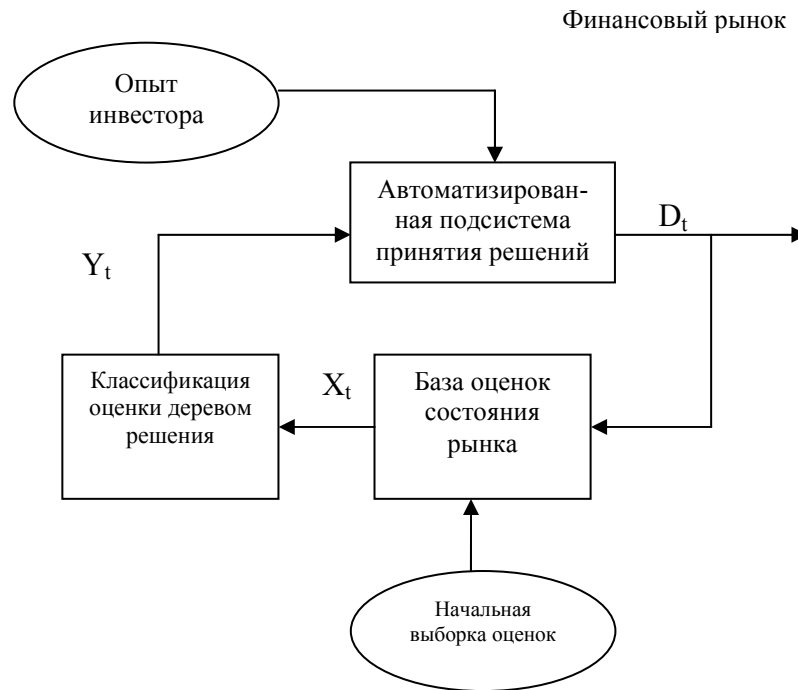


Рис. 2. Структурная схема автоматизированной системы принятия решений

С помощью начальной выборки оценок состояния рынка система формирует дерево решений, которое в дальнейшем позволит классифицировать последующие векторы оценок в класс роста прибыли. Получив результаты классификации, система производит совет инвестору, на основе которого и учитывая собственный опыт инвестор принимает решение, что заносится в базу оценок состояния рынка и в дальнейшем будет использоваться системой для классификации следующей оценки.

### Выводы

Разработана методология повышения эффективности принятия решений на основе классификации оценок состояния рынка с помощью метода дерева решений. Для этого была составлена структурная схема автоматизированной экспертной системы, способной к самообучению и выработке совета инвестору для работы на финансовом рынке. Кроме того, авторами была построена модель принятия решений, с помощью которой система принимает решение, описаны ее характеристики и описана принцип работы. Следующим шагом для разработки экспертной системы и получения практических результатов авторами планируется сформулировать алгоритм синтеза дерева решений и алгоритм сокращения его узлов для повышения точности классификации, проверить работу системы в реальных условиях работы на финансовом рынке.

### СПИСОК ЛИТЕРАТУРЫ

1. Попов Е.П., Бесекерский В.А. Теория систем автоматического регулирования. – М.: Наука, 2003. – 759 с.
2. Берестнева О.Г., Муратова Е.А. Построение логических моделей с использованием деревьев решений // Известия Томского политехнического университета. – 2004. – Т. 207, вып. 2. – С. 55-61
3. J. R. Quinlan, Improved Use of Continuous Attributes in C4.5 // Journal of Artificial Intelligence Research 4. – 1996. P. 77-90.

**Кветный Роман Наумович** – д.т.н., проф., заведующий кафедры автоматики и информационно-измерительной техники.

**Коцюбинський Владимир Юрьевич** – к.т.н, доцент кафедры автоматики и информационно-измерительной техники.

**Козачко Алексей Николаевич** – к.т.н., ассистент кафедры инженерной и компьютерной графики.

**Кислица Людмила Николаевна** – магистр, соискатель кафедры автоматики и информационно-измерительной техники, [lus83@mail.ru](mailto:lus83@mail.ru).

**Казимирова Нина Владимировна** – студентка кафедры автоматики и информационно-измерительной техники, [nkazimirova@spilnasprava.vn.ua](mailto:nkazimirova@spilnasprava.vn.ua).

Винницкий национальный технический университет