На завершення сформулюємо основні переваги ІІС з кількома інтерпретаторами порівняно з інтелектуальними системами, що містять одну МЛВ:

- прискорення часу розв'язання задачі;
- можливість використання різних функцій керування залежно від підзадачі;
- використання в одній системі кількох методів представлення знань;
- перевірка роботи системи заміною одних інтерпретаторів на інші;
- збільшення надійності отриманого розв'язку;
- координація роботи системи загалом.

1. Буч Г. Объектно-ориентированный анализ и проектирование. - М.: Издательство Бином, СПб: Невский диалект, 1998. 2. Клыков Ю.И., Горьков Л.Н. Банки данных для принятия решений. - М.: Советское радио, 1980. 3. Courtois, P. On time and Space Decomposition of Complex Structures. Communications of the ACM vol. 28(6), - 1985. - p. 596. 4. Моисеев Н.Н. Математические задачи системного анализа. - М.: Наука, 1981. 5. Simon, H. The Sciences of the Artificial. Cambridge, MA: The MIT Press, 1982. - p. 218 6. Литвин В.В. Формальна модель інтелектуальної інформаційної системи // Вісн. ДУ "Львівська політехніка". - 1999. - № 386. - С. 104-108. 7. Дж. Ортега "Введение в паралельные и векторные методы решения линейных систем". - М.: Мир, 1991. 8. Литвин В.В., Нікольський Ю.В. Застосування методів логічного програмування для автоматизації процедур прийняття рішень у діяльності служби працевлаштування // Вісн. ДУ "Львівська політехніка". - 1998. - № 330 - С. 153-163.

УДК 683.1

Е. Марецька*, Р.Б. Кравець
* Інститут інформатики та управління,
НУ "Львівська політехніка", кафедра "Інформаційні системи та мережі"

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ У СФЕРІ СОЦІАЛЬНОГО СТРАХУВАННЯ

© Марецька Е., Кравець Р., 2002

This paper considers classification method for knowledge discovery in databases. The induction of decision tree algorithms for supervised classification is described. Authors propose to use this method to predict the result of sport match.

Стаття містить класифікаційні методи видобування знань у базах даних. Описаний метод індукції у деревах рішень. Автори пропонують використовувати ці методи для передбачення результату спортивних матчів

1. Вступ

Інформаційні технології все глибше проникають у різні сфери економічної діяльності. Однією з таких галузей є соціальне страхування. Для ефективного ведення страхового бізнесу доводиться фіксувати і аналізувати велику кількість параметрів, які впливають на виникнення страхових випадків. При цьому страховим компаніям доводиться враховувати усі ці фактори і приймати рішення, спрямовані на покращання якості виконання обов'язків перед клієнтами і водночас на отримання максимального прибутку.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Авторами розглядається математична модель соціального страхування, метод видобування знань на основі класифікації даних, а також пропонується застосування цього методу для аналізу настання страхових випадків.

2. Математична модель соціального страхування

Математична модель страхування життя базується на таблицях тривалості життя. Ці таблиці у Польщі складаються Головним статистичним управлінням кожні п'ять років окремо для чоловіків та жінок з міської та сільської місцевості. У таблицях містяться умовні імовірності $p(1/n)$ прожиття одного року особою, яка має вік n років ($n = 0,\ldots,99$).

Виходячи з умовних імовірностей $p(1/n)$, можна обчислити:

- ймовірність того, що особа житиме у віці n років (тобто заплатить страховий внесок):

$$P_n = P_{n-1} \cdot p\left(1/_{n-1}\right), \quad (1)$$

- ймовірність того, що особа помре у віці n років (тобто буде виплачена страховка):

$$Q_n = P_{n-1} \cdot q\left(1/_{n-1}\right). \quad (2)$$

У формулах (1) та (2) прийнято:

$$P_0 = 1,$$

$$q\left(1/_{n-1}\right) = 1 - p\left(1/_{n-1}\right).$$

Імовірності P_n та Q_n відіграють ключову роль в обчисленнях внесків та страхових виплат. Очікуваний прибуток страхового товариства (надходження страхових внесків) повинен зрівноважуватись очікуваними витратами (страховими виплатами), тобто сума ймовірних і дисконтованих страхових внесків дорівнює сумі ймовірних і дисконтованих страхових виплат. Математичний запис цієї умови виглядає так:

$$\sum P_n \cdot S_n / D_n = \sum Q_n \cdot W_n / D_n, \quad (3)$$

де S_n - страховий внесок у n-му періоді, W_n - страхова виплата у n-му періоді, D_n - дисконтний показник з n-го періоду на період 0.

Обчислення страховок є двоїстою задачею: для заданих страхових внесків можна визначити страхові виплати – або для заданих страхових виплат можна визначити страхові внески.

Дисконтний показник D_n залежить від стану ринку і може описуватись так:

- просте дисконтування зі сталим відсотком

$$D_n = 1 + r \cdot n \quad (4)$$

- просте дисконтування із змінним відсотком

$$D_n = 1 + r_1 + \ldots + r_n \quad (5)$$

- складне дисконтування зі сталим відсотком

$$D_n = (1 + r)^n \quad (6)$$

- складне дисконтування із змінним відсотком

$$D_n = (1 + r_1) \cdot \ldots \cdot (1 + r_n), \quad (7)$$

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

де r_n - відсоток за n-й період.

Кожна модель страхування життя може бути подана у різних станах ринку.

2.1. Страхування приданого

При страхуванні приданого одна особа сплачує страхові внески, а інша може отримати страхові виплати як придане.

Обчислення страховки відбувається з дотриманням умови рівності капіталів:

$$\sum P_n^1 \cdot S_n / D_n = P_N^2 \cdot W_N / D_N , \tag{8}$$

де P_n^1 - імовірність сплати чергового страхового внеску, P_N^2 - імовірність виплати приданого.

Страхова угода передбачає, що страхове товариство виплатить придане W_N навіть якщо особа, що сплачує внески, помре і не сплатить усіх N внесків. З іншого боку навіть при сплаті усіх внесків виплата приданого не відбудеться, якщо особа, що повинна його отримати, не виконає умов страхування в термін N .

2.2. Страхування життя

Страхування життя на термін N може мати два варіанти:

- з виплатою у випадку смерті в термін $n \le N$;
- з виплатою у випадку доживання до терміну N .

Обчислення страховки у першому варіанті базується на виконанні умови рівності капіталу (3). У другому випадку сума страховки обчислюється на основі формули рівності капіталу (8).

У загальному випадку страхування життя охоплює обидва варіанти одночасно. Тоді для обчислення суми страховки до правої частини рівняння (3) слід додати праву частину рівняння (8).

Загальне страхування життя дозволить завжди отримати страхову виплату.

2.3. Страхування життя з інвестиційним фондом

Страхування життя з інвестиційним фондом полягає у поділі страхового внеску на частини: страхову U_n та інвестиційну I_n .

$$S_n = U_n + I_n \tag{9}$$

На основі внесків U_n отримуються виплати W_n відповідно до умови рівності капіталу (3). Внески в інвестиційний фонд приносять окремі виплати:

$$V_m = \sum I_n \cdot O_n , \tag{10}$$

де O_n - показник процентування внесків I_n .

Показники процентування залежать від стану ринку:

- просте процентування зі сталим відсотком

$$O_n = 1 + r \cdot (N - n) \tag{11}$$

- просте процентування із змінним відсотком

$$O_n = 1 + r_{n+1} + \ldots + r_N \tag{12}$$

- складне процентування зі сталим відсотком

$$O_n = (1 + r)^{N-n} \tag{13}$$

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

- складне процентування із змінним відсотком

$$O_n = (1 + r_{n+1}) \cdot \ldots \cdot (1 + r_N) \tag{14}$$

У випадку смерті у термін $m \le N$ виплачуються дві частки W_m замість V_m. При доживанні до віку N виплачується тільки частка V_N. Відповідний поділ внеску S_N дозволить отримати потрібну виплату $W_m + V_m$.

2.4. Пенсійне страхування

Пенсійне страхування полягає у сплаті M внесків з імовірностями P_m, $m = 0, \ldots, M-1$ до певного терміну, а потім отриманні N виплат з імовірностями P_n, $n = 1, \ldots, N$ з нагромадженого фонду. Розмір накопиченого фонду F_0 обчислюється за формулою

$$F_0 = \sum P_m \cdot S_m \cdot O_m . \tag{15}$$

Чергова виплата за кожен період $1, \ldots, N$ відбувається з імовірністю P_n. Розмір цих виплат обчислюється на основі правої частини рівняння (3), тобто

$$F_0 = \sum P_n \cdot W_n / D_n \tag{16}$$

У процесі здійснення виплат у періоди $n = 1, \ldots, N$ розмір пенсійного фонду F_n зменшується, так що $F_N = 0$.

2.5. Страхування кредитів

Перед наданням кредиту банк вимагає забезпечення повернення цього кредиту. Забезпеченням повернення кредиту від фірми є її майно. У випадку фізичних осіб, яким потрібен кредит для ведення підприємницької діяльності, купівлі нерухомості або споживчих товарів (комп'ютер, камера і т.п.) – забезпеченням повернення кредиту може бути страховий поліс.

Вирізняють такі кредити, повернення яких може забезпечуватись страховим полісом фізичної особи:

- споживчий кредит
- іпотечний кредит
- інвестиційний кредит

2.5.1. Споживчий кредит

У випадку споживчого кредиту страховий поліс дозволить сплатити суму кредиту, яка не була б сплачена у разі смерті чи хвороби боржника. Оскільки виплати по страховому полісу з часом зменшуються, тому що зменшуються сума боргу. Страховий внесок у загальному є одноразовий, і сплачується в момент надання споживчого кредиту. Внесок повинен зрівноважувати ймовірні і дисконтовані виплати на термін $n = 0$.

Математична модель складається з трьох етапів.

1. Визначення виплат по кредиту

З умови рівності капіталу отримаємо:

$$K_0 \cdot O_0 = \sum R_n \cdot O_n \tag{17}$$

Далі для визначення R_n маємо:

$$R_n = R_1 + (n-1) \cdot \Delta R, n = 2, \ldots, N \tag{18}$$

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

де K_0 - сума кредиту, R_0 - кінцева сума виплати в n-му періоді, ΔR - показник нарощування вартості виплат по кредиту, O_n - показник процентування.

2. Визначення страхових виплат

З умови рівності капіталу отримуємо виплати як суми дисконтованих несплачених виплат по кредиту:

$$W_m = \sum R_n / D_n \text{ для } n \geq m, m = 1, \ldots, N \quad (19)$$

3. Визначення страхових внесків

Маючи виплати W_m, імовірності P_m та Q_m та дисконтні показники D_m, можна з умови рівності капіталу (3) визначити суми внесків S_m. Суми внесків можуть збільшуватись за рахунок нарощування їх вартості.

2.5.2. Іпотечний кредит

Як правило, іпотечним кредитом є довготривалий кредит на купівлю нерухомості для сім'ї. Сім'я повинна мати можливість проживання в купленому будинку навіть у випадку смерті, хвороби чи втрати працездатності особи, що взяла кредит. У такому випадку кредит сплачується в опроцентованих частках.

$$I_n = K_0 \cdot r_n, n = 1, \ldots, N \quad (20)$$

Капітал не сплачується загалом. Після закінчення виплати по відсотках йде наступний аналогічний кредит на виплату капіталу.

Сума виплати по страховому полісу W_n є величиною сталою і дорівнює вартості капіталу кредиту K_0.

У випадку смерті страхувальник виплачує банку частку капіталу K_0, а сума боргу по відсотках обнулюється. Величина відсотків, обчислена за формулою (20), задовольняє умові рівності капіталу.

Страховий внесок може ділитися на частини. Величина цих часток обчислюється, виходячи з умови рівності капіталу (3).

2.5.3. Інвестиційний кредит

У випадку інвестиційного кредиту, наданого фізичній особі, яка провадить підприємницьку діяльність, обумовлюється, що кредит сплачується комбінованими внесками (по капіталу та по відсотках). Кредит повинен бути виплачений повністю в термін N. Забезпеченням сплати цього кредиту є страхування життя на термін N. Виплати по страховому полісу з часом зменшуються. Страховий внесок може ділитися на частини.

Математична модель страхування інвестиційного кредиту аналогічна до моделі споживчого кредиту з тією різницею, що слід окремо розглядати внески по капіталу та по відсотках.

2.6. Узагальнення моделі

В умову рівності капіталу входить N внесків S_n та N виплат W_n. З одного рівняння не можна визначити N невідомих. З огляду на те приймається, що внески та виплати утворюють арифметичні або геометричні прогресії, тобто:

$$S_n = S_0 + n \cdot \delta S \quad (21)$$

$$S_n = S_0 \cdot (1 + i)^n \quad (22)$$

де δS - сума приросту внесків, i - ступінь індексації внесків.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Аналогічні рівняння можна прийняти і для виплат.

Перераховані вище види страхування розглядалися для однієї особи. Загальнішою моделлю є страхування групи осіб (наприклад, сім'ї з двох осіб). У такому випадку для розрахунків внесків та виплат потрібно визначити ймовірність Q_n смерті групи осіб. Наприклад, можна вважати, що група існує (сплачує внески або отримує страховку), коли живе принаймні один член цієї групи (альтернатива). Так само можливий випадок, коли група існує (платить внески або отримує виплати), поки живі усі її члени.

Відповідні ймовірності для моделі групового страхування визначаються наступними співвідношеннями.

Кон'юнкція:

$$P_n = P_n^1 \cdot P_n^2 \tag{23}$$

Альтернатива:

$$P_n = P_n^1 \cdot Q_n^2 + Q_n^1 \cdot P_n^2 \tag{24}$$

3. Метод класифікації даних

Метод класифікації даних складається з двох основних етапів [7]. На першому етапі здійснюється аналіз даних, що зберігаються в базі даних, і будується модель, яка кожному об'єкту ставить у відповідність мітку класу, до якого цей об'єкт належить. Сукупність об'єктів, по яких створюється класифікаційна модель, називається *навчальним набором*.

Побудована модель подається деяким способом представлення знань, переважно у формі дерев рішень або класифікаційних правил. Класифікатор може мати як детермінований характер, коли кожному об'єкту ставиться у відповідність точно один клас, так і недетермінований, коли об'єкт належить до кількох класів з певним розподілом деяких мір невизначеності (наприклад, підтримки, ступеня довіри тощо).

На другому етапі створена класифікаційна модель використовується для класифікації нових об'єктів. Сукупність цих об'єктів називається *тестовим набором даних*.

Розглянемо формальну постановку задачі класифікації даних. Нехай маємо навчальний набір даних L, який складається з кортежів $t \in L$, а також множину класів C, яка складається з міток класів $c_i \in C$, $i = 1,\ldots,m$. Для кожного кортежу з навчального набору відомо, до якого класу він належить, тобто кортежі мають структуру $\langle A, c \rangle$, де A – атрибути, що описують об'єкт, c – мітка класу. Крім того, задано тестовий набір даних T, для кортежів якого не відомо, до якого класу вони належать. Потрібно (1) на основі навчального набору побудувати відображення $K: L \to C$ у формі знань, яке кожному кортежу t навчального набору L ставить у відповідність мітку c з множини класів C; (2) кожному кортежу t тестового набору T, використовуючи відображення K, поставити у відповідність мітку c з множини класів C.

3.1. Метод класифікації на основі індукції дерев рішень

Побудова класифікаційної моделі здійснюється у такій послідовності [3].

3.1.1. Вибірка даних

На першому кроці здійснюється вибірка даних із бази даних, які становитимуть навчальний набір. При цьому особа, яка приймає рішення, чи експерт аналізує предметну область та формує множину параметрів, які описують об'єкт дослідження і можуть впливати на вихідне рішення. У результаті вибірки створюється відношення, яке називається

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

таблицею рішень, зі структурою типу *R(A,D)*, де *A* – набір атрибутів-факторів, значення яких впливають на формування рішення, *D* – набір атрибутів рішення.

У соціальному страхуванні таблиця рішень містить дані про фактори, які впливають на настання страхового випадку, а як атрибут рішення виступає вид страхового випадку.

3.1.2. Попереднє опрацювання даних

На цьому кроці здійснюється поповнення даних з невизначеностями, опрацювання аномальних даних, дискретизація числових величин та виділення із загального набору параметрів підмножини суттєвих факторів.

Після вибірки даних частина кортежів таблиці рішень може містити невизначені або пропущені значення. Крім того, деякі значення певного атрибута можуть різко відрізнятися від решти значень. Для коректного виконання алгоритму побудови дерева рішень усі невизначеності та аномалії потрібно або усунути з таблиці рішень, або довизначити і згладити їх на основі статистичного аналізу значень атрибута [1].

Для зменшення кількості можливих значень числових атрибутів та побудови дерева рішень на загальнішому рівні агрегації даних здійснюється дискретизація числових величин [7]. При цьому числовий домен атрибута розбивається на сукупність інтервалів, і кожне значення атрибута у таблиці рішень заміняється на відповідний інтервал.

Одним із критеріїв якості отриманих знань є їх цілковите розуміння особою, що приймає рішення [3]. Тому важливо, щоб дерево рішень мало достатньо просту структуру. Одним із способів досягнення цього є виділення серед атрибутів-характеристик тих, що найбільше впливають на вихідне рішення. Цей крок здійснюється за допомогою методів факторного аналізу [2].

3.1.3.Побудова дерева рішень та набору класифікаційних правил

На цьому кроці виконується алгоритм ID3 [14] побудови дерева рішень. Алгоритм виконується у такій послідовності:

1. Створюється початковий вузол дерева
2. Якщо усі кортежі навчального набору належать до одного класу, то вузол визначається як листковий, і йому присвоюється мітка класу
3. Інакше алгоритм використовує *інформаційний приріст* [14] для визначення атрибута, на основі якого добудовується дерево
4. Для кожного значення вибраного атрибута формується гілка дерева, і кортежі навчального набору, що залишилися, розділяються на відповідні підмножини
5. Алгоритм виконується рекурсивно для побудови піддерева на основі набору атрибутів, що залишилися; при цьому атрибути, які вже були задіяні при побудові дерева, не розглядаються
6. Алгоритм зупиняється при настанні однієї з таких умов:
    - Усі кортежі навчального набору, що залишилися, належать до одного класу
    - Для побудови дерева рішень використані усі атрибути
    - Для побудови дерева використані усі кортежі навчального набору

На основі побудованого дерева рішень формується набір класифікаційних правил типу “ЯКЩО <атрибут>=<значення> ТО <страховий випадок>=<значення>”.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Описаний тут підхід дозволяє страховій компанії визначати вплив різноманітних факторів на виникнення страхових випадків, і приймати рішення про надання послуг соціального страхування.

1. Айвазян С.А. и др. Прикладная статистика: Основы моделирования и первичная обработка данных. – М.: Финансы и статистика, 1983. 2. Иберла К. Факторный анализ. – М: Статистика, 1980. 3. Advances in knowledge discovery and data mining. Fayyad U.M., Piatetsky-Shapiro G., Smyth P., Uthurusamy R. (editors), AAAI/MIT Press, 1996. 4. Bijak W., Podgórska M., Utkin J.M., Nowak E. Matematyka I statystyka finansowa. //Warszawa, FRR, 1994. 5. Cyrek M. Komputerowy system ubezpieczeń. //Bielsko-Biała, WZIIZ, 1998, (praca dyplomowa inżynierska). 6. Dobija M., Smaga E. Podstawy matematyki finansowej I ubezpieczeniowej. Warszawa-Kraków, WN PWN,1996. 7. Han J., Kamber M. Data mining: methods and technique, Morgan Kaufman, 2000. 8. Jelińska J. Komputerowa analiza ubezpieczeń kredytów konsumpcyjnych. //Bielsko-Biała, WZIIZ, 1999, (praca dyplomowa inżynierska). 9. Klenart M. Komputerowa analiza ubezpieczeń z funduszem inwestycyjnym. //Bielsko-Biała, WZIIZ, 1999, (praca dyplomowa inżynierska). 10. Kohut I. Komputerowa analiza ubezpieczeń kredytów hipotecznych. //Bielsko-Biała, WZIIZ, 1999, (praca dyplomowa inżynierska). 11. Kózka A. Komputerowa analiza ubezpieczeń emerytalnych. //Bielsko-Biała, WZIIZ, 1999, (praca dyplomowa inżynierska). 12. Matematyka finansowa. Warszawa, Bizant, 1994. 13. Przeradzki K. Komputerowa analiza ubezpieczeń posagowych. //Bielsko-Biała, WZIIZ, 1998, (praca dyplomowa inżynierska). 14. Quinlan J.R. Induction of decision trees. Machine Learning, 1:81-106, 1986.

УДК 683.1

А.М. Пелещишин
НУ “Львівська політехніка”,
кафедра “Інформаційні системи та мережі”

ОПТИМІЗАЦІЯ РОЗМІЩЕННЯ ІНФОРМАЦІЙНИХ РЕСУРСІВ НА КОМПОНЕНТАХ WEB-СИСТЕМИ

© Пелещишин А.М., 2001

This paper includes the research of the Web-system coordination problem. proposed. Typical example of Web-system coordination are investigated. This example is optimization of query support location of global Web-system.

У статті розглядається проблема координації складної Web-системи. Досліджено типовий приклад координації - задачу оптимізації розміщення інформаційних ресурсів на компонентах Web-системи

Задача побудови інформаційних систем оптимальної фізичної структури досліджувалася (зокрема, у роботах [4,9,10,13,16,17]), проте з розвитком глобальної мережі Інтернет та Web-технологій виникли принципово нові види цієї задачі. У нових умовах побудова систем оптимальної структури здійснюється з врахуванням глобального, багаторівневого характеру ІС, відсутності єдиної власності на них, стохастичної поведінки користувачів такої системи. Ефективним математичним апаратом розв’язання таких задач є (як показано в роботах [8,2,14]) теорія координації ([1,5,6]).

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua