

ПІДГОТОВКА БАЗИ ЗНАНЬ БАЙЕСОВСЬКОЇ ЕКСПЕРТНОЇ СИСТЕМИ В УМОВАХ ЗНАЧНОЇ КОРЕЛЯЦІЇ РЕЗУЛЬТАТІВ ДІАГНОСТИЧНИХ ЕКСПЕРИМЕНТІВ

Анатолій Власюк, Віталій Гнатюк, Руслан Белзецький

Вінницький національний технічний університет

Хмельницьке шосе, 95, Вінниця, 21021, Україна, тел.: (0432) 43-52-60, E-Mail: vit@vstu.vinnica.ua

Анотація

У доповіді розглянуто принципи розробки інформаційних систем, для обробки великого об'єму інформації, яку практикуючому лікарю необхідно постійно отримувати та аналізувати для оптимального вибору підходів до діагностування та лікування кожного пацієнта. Показано, що в таких системах доцільно використовувати розроблену авторами експертну систему. Запропоновано методу підготовки бази знань для цієї системи, яка на відміну від відомих методик кластерного та факторного аналізу формує ієрархічну структуру бази знань, придатну для ефективної роботи експертної системи, в умовах коли багато експериментів мають сильну кореляцію властивостей, тобто дають залежні свідчення, та можливе суттєве спотворення діагностичної інформації.

Постановка проблеми. Мета дослідження

Стрімке розширення номенклатури методів лікування, асортименту лікарських засобів та біологічно активних добавок призводить до того, що практикуючому лікарю дуже важко бути обізнаним у можливостях і особливостях продукції, яку пропонують різні виробники. Ситуація погіршується, при порівнянні препаратів з однієї якісної та цінової категорії різних виробників для того щоб вибрати оптимальний для даної людини комплекс препаратів або схему лікування. У цьому випадку вибір як правило робиться не з повної множини аналогів, а серед декількох препаратів або схем лікування добре відомих лікарю. Тому такий вибір може виявитися не досить вдалим і, в порівнянні з оптимальним вибором, призвести до втрати часу до початку дійсно ефективного в даному випадку лікування або призвести до отримання суттєво більших небажаних побічних ефектів.

З іншої сторони, виробники, як правило, використовують інформаційні технології для просування на ринок своїх методів лікування, лікарських засобів та біологічно активних добавок, і відкриті для активної співпраці з посередниками на інформаційному ринку та готові надавати актуальну інформацію про свої розробки в необхідному вигляді. Це дозволяє створити та підтримувати в актуальному стані експертну систему (ЕС) орієнтовану на практикуючих лікарів. Особливістю розробленої нами ЕС, є динамічне компонування бази знань (БЗ) із предметних баз знань (ПБЗ) незалежно розроблених виробниками, які використовуються в процесі діагностування та прийняття рішень щодо лікування. Це реалізовано у вигляді віртуальної системи діагностики та корекції стану організму людини [1], яка створює для лікаря можливість врахування всього наявного об'єму інформації про препарати та можливість ефективно застосовувати без залучення відповідних фахівців різноманітні апаратні методи визначення функціонального стану здоров'я людини, які базуються на вимірюванні параметрів біологічно активних точок та на використанні інших переносних датчиків (електрограми, пульсодіагностика, оптичні, ...).

Указана ЕС використовує запропановану нами модифікацію байєсовського методу логічного виведення [2, 3] та ієрархічну структуру БЗ, основу на фреймах. Така організація логічного виведення ставить вимоги до БЗ, які мають виконуватися у межах кожного класу. Основні з них:

- можливість бути доведеною або відхиленою для кожної гіпотези, тобто перекриття неточностей підготовки БЗ та можливих похибок у діагностичній інформації надлишковістю свідочств;
- незалежність свідочств або компенсація відповідної похибки рівномірним розподілом надлишковості свідочств;
- нормалізація параметрів фреймів, що представляють різні ПБЗ.

Готуючи різні БЗ для тестування процесу логічного виведення, автори з'ясували, що структурування БЗ методами факторного або кластерного аналізу (їх огляд наведено у [4]) часто утворює класи, порушуючи наведені вимоги до БЗ. Тому виникла потреба доповнити указані методи діями, що забезпечать утворення лише таких класів, для яких виконуються наведені вимоги до БЗ.

Вирішення проблеми

Дотриматись цих вимог при формуванні БЗ можливо за допомогою наступної процедури, запропонованої авторами.

На початковому етапі формування БЗ включаємо до неї всі можливі експерименти не зважаючи на дублювання інформації, яку вони дозволяють отримати та виникаюче через це розбалансування БЗ.

Перший етап структурування БЗ: Виходячи з обраного методу кластерного аналізу визначити міри таксонометричної подібності між всіма гіпотезами.

Другий етап — побудова попереднього дерева гіпотез за допомогою наступного ітераційного процесу. Спочатку об'єднуємо в один клас гіпотези, які відповідно критеріям обраного методу кластерного аналізу (автоматичної класифікації) є першими претендентами на об'єднання. Таким чином утворюється перше піддерево. Замість формально обраного першого претендента, можна використати іншу, з пріоритетних відповідно критеріям обраного методу кластерного аналізу, групу претендентів на об'єднання в клас, семантично ближчу до звичної для майбутнього користувача класифікації та більш зрозумілу для нього. Далі переходимо до розгляду утвореного класу замість гіпотез, що увійшли в нього. Визначаємо міри таксонометричної подібності нової гіпотези, відповідної утвореному класу, та решти гіпотез, що розглядаються. Ці шаги ітерації повторюємо доки всі гіпотези не будуть об'єднані в одне попереднє дерево гіпотез.

На протязі даного етапу структурування БЗ після утворення кожного нового класу перевіряємо його на відповідність наведеним раніше вимогам до БЗ. Якщо ці вимоги не виконуються доповнюємо БЗ експериментами які вирішують проблему, що виникла. На цьому етапі зайві експерименти не вилучаються з БЗ, навідь якщо це вилучення могло відновити виконання вимог до БЗ.

Третій етап — перетворення попереднього дерева гіпотез у БЗ ЕС. Для ефективного використання ЕС доцільно об'єднувати в клас таку кількість гіпотез, щоб процес вибору між ними, який робить ЕС, залишався прозорим для користувача, тобто до 10...30. При цьому бажано робити таке укрупнення класів відповідно до ієрархії, яка природньо властива предметній галузі. Цей процес укрупнення класів шляхом розкриття недоцільної інкапсуляції ефективніше проводити від верхнього рівня ієрархії в глибину.

На протязі даного етапу структурування БЗ, після утворення кожного нового класу, перевіряємо його на відповідність наведеним раніше вимогам до БЗ. Якщо ці вимоги не виконуються, то корегуємо набір експериментів, пов'язаних з даним класом гіпотез. На цьому етапі, виконання наведених вимог до БЗ не є суворо обов'язковим для кожного рівня ієрархії та для кожного класу, але є бажаним так як забезпечує ефективність роботи ЕС.

Запропонований алгоритм автоматизованої класифікації пов'язує з гіпотезою, що представляє собою клас, тільки ті експерименти, які пов'язані з усіма гіпотезами, що увійшли в цей клас. Тому, якщо це не погіршує прозорість семантики побудови БЗ, укрупнення класів бажано проводити таким чином, щоб при цьому покращувалась надійність (якість) розпізнавання гіпотез, які входять в клас, що формується, в порівнянні з гіпотезами, які входили в клас, що був розкритий при даному укрупненні.

Методи факторного аналізу можуть бути доповнені аналогічно.

Висновки

Запропонована методика класифікації гіпотез у БЗ, на відміну від відомих методик кластерного та факторного аналізу, здатна формувати БЗ із ієрархічною структурою, придатну для ефективної роботи ЕС, в умовах коли багато експериментів мають сильну кореляцію властивостей, тобто дають залежні свідчення, що характерно для задач медичної діагностики та задач підбору медичних препаратів.

Вона може бути використана при створенні віртуальних розподілених інтелектуальних інформаційно-вимірювальних систем основаних на ЕС, які використовуються для прийняття рішень на основі діагностування або інтерпретації спостережень в різних предметних областях.

Література:

- [1] Власюк А., Гнатюк В., Власюк Н. Інформаційне забезпечення віртуальної системи діагностики та корекції стану організму людини // Інтернет-Освіта-Наука-2006, п'ята міжнародна конференція ІОН-2006. 10—14 жовтня, 2006. Збірник матеріалів конференції. Том 2. — Вінниця: Універсум-Вінниця, 2006. — С. 574—576
- [2] Роїк О. М., Гнатюк В. В., Власюк Н. А. Ієрархічне структурування бази знань медичної системи телемоніторингу стану організму людини // Вісник Вінницького політехнічного інституту. — 2007 — №3. — С. 78—82.
- [3] Власюк А. І., Гнатюк В. В., Белзецький Р. С. Ієрархічне структурування бази знань експертної системи з байєсовським методом логічного виведення // Optoelectronic Information Technologies «Photonics-ODS-2008». Збірник тез доповідей IV Міжнародної науково-технічної конференції, м. Вінниця, 30 вересня—2 жовтня 2008 року. — Вінниця: Універсум-Вінниця, 2008. — С. 137—138
- [4] Электронный учебник StatSoft // StatSoft inc. — 2001. Режим доступа до кн. : <http://www.statsoft.ru/home/textbook/default.htm>