**Терентьєв Олександр Миколайович. Моделі і методи побудови та аналізу байєсівських мереж для інтелектуального аналізу даних : Дис... канд. наук: 05.13.06 - 2009.**

|  |  |
| --- | --- |
| |  | | --- | | Терентьєв О.М. Моделі і методи побудови та аналізу байєсівських мереж для інтелектуального аналізу даних. – Рукопис.  Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Національний технічний університет України “Київський політехнічний інститут”, Київ, 2009 р.  Робота присвячена проблемі підвищення швидкості та якості інтелектуального аналізу даних шляхом розробки нової методики побудови та застосування дискретних байєсівських мереж (БМ), методу формування ймовірнісного висновку і створення на їх основі нової системи підтримки прийняття рішень. Зроблено огляд і аналіз методів інтелектуального аналізу даних та обґрунтована ефективність застосування апарату БМ.  Для побудови структури БМ запропоновано евристичний метод навчання лінійної складності за статистичними даними. Для визначення міри зв’язку між вершинами метод використовує значення взаємної інформації, а в якості функції оцінювання структури моделі – опис мінімальною довжиною (ОМД). Також розроблено метод побудови точного ймовірнісного висновку в БМ за навчальними даними. Для обчислення значень ймовірностей станів вершин замість таблиць умовних ймовірностей використовується матриця емпіричних значень спільного розподілу ймовірностей всієї мережі. Для оцінки якості побудованої структури отримано нову модифіковану функцію, яка на відміну від функції Купера-Герсковича, суттєво прискорює процес обчислення та знімає обмеження на розмір вибірок навчальних даних.  Розроблена і програмно реалізована оригінальна система підтримки прийняття рішень для інтелектуального аналізу даних на основі БМ, яка грунтується на запропонованих методах побудови структури та ймовірнісного висновку. СППР впроваджена в першій київській філії ВАТ VAB “ВіЕйБі Банк”, що дало можливість побудувати ефективні прогнозуючі моделі для підтримки прийняття рішень з метою оцінки ризиків при кредитуванні фізичних осіб. Наукові результати використані у навчальних курсах НТУУ „КПІ”. | |
| |  | | --- | | 1. Виконано аналіз методів інтелектуального аналізу даних та обґрунтована ефективність застосування апарату байєсівських мереж (БМ). Встановлено, що основним недоліком методів побудови топології БМ є нелінійна поліноміальна складність обчислень; недоліком існуючих методів формування ймовірнісного висновку є їх обчислювальна складність, яка залежить від кількості вершин мережі та кількості дуг, які з’єднують ці вершини мережі між собою, а також неоднозначність – різні методи можуть давати різні числові результати.  2. Для побудови структури БМ запропоновано евристичний метод навчання лінійної складності за статистичними даними. Для визначення міри зв’язку між вершинами метод використовує значення взаємної інформації, а в якості функції оцінювання структури моделі – опис мінімальною довжиною (ОМД). Швидкодія побудови моделей, у порівнянні з методами повного перебору, зростає в 10-20 разів.  3. Розроблено метод побудови точного ймовірнісного висновку в БМ за навчальними даними. Для обчислення значень ймовірностей станів вершин замість таблиць умовних ймовірностей використовується матриця емпіричних значень спільного розподілу ймовірностей всієї мережі. Головними перевагами методу є залежність швидкості роботи тільки від розміру навчальної вибірки, відсутність потреби у попередньому перетворенні структури БМ та простота її реалізації. Швидкодія побудови ймовірнісного висновку зростає в 2-5 разів.  4. Для оцінювання якості побудованої структури запропоновано нову модифіковану функцію, яка, на відміну від функції Купера-Герсковича, суттєво прискорює процес обчислень та знімає обмеження на розмір вибірок навчальних даних.  5. Розроблена і програмно реалізована оригінальна інформаційна система підтримки прийняття рішень для інтелектуального аналізу даних на основі БМ, яка грунтується на запропонованих методах побудови структури та ймовірнісного висновку. СППР впроваджена в першій київській філії ВАТ VAB “ВіЕйБі Банк”, що дало можливість побудувати ефективні прогнозуючі моделі для підтримки прийняття рішень з метою оцінювання ризиків кредитування фізичних осіб.  Наукові результати використано у навчальних курсах з математичного моделювання і проектування комп’ютерних інформаційних систем. Рекомендується подальше використання отриманих результатів при розв’язуванні практичних системних задач різної природи та розробці нових методів інтелектуального аналізу даних на основі мереж Байєса. | |