

НАВЧАЛЬНА ПРОГРАМА

дисципліни «Інтелектуальна обробка даних»

Підготовки магістрів
За спеціальністю 8.05010103- Системне проектування
Напряму підготовки 6.050101 – Комп'ютерні науки

РОЗГЛЯНУТО

На засіданні кафедри «Системи і процеси управління»

Протокол № 9
» 21 06 2016р

Завідуючий кафедрою
проф. Бреславський Д.В.
(вчене звання, прізвище, ініціали)

ЗАТВЕРДЖЕНО

Вченою радою факультету
_____ інженерно-фізичного

від «Протокол № _____» _____ 2016_р

Декан факультету
проф. Конкін В.М.
(вчене звання, прізвище, ініціали)



Харків, 2016

Інтелектуальні інформаційні системи

Мета та завдання навчальної дисципліни засвоєння студентами основних теоретичних відомостей та практичних вмінь з курсу. Підготувати студента до ефективного використання як класичних, так і сучасних методів інтелектуального аналізу даних та обробки інформації з використанням м'яких обчислень та нейромереж для проектування елементів автоматизованих комп'ютерних систем та їх використання у подальшій професійній діяльності; допомогти набути навички практичної роботи з програмними засобами для інтелектуальної обробки даних.

Завдання дисципліни – оволодіння студентами науковими основами, сучасною методологією та особливостями застосування інтелектуальної обробки даних; засвоєння майбутніми фахівцями теоретичних основ інформаційних систем, орієнтованих на застосування стандартів Data Mining; набуття уміння програмувати окремі елементи систем Data Mining різного призначення і різної проблемної орієнтації на всіх стадіях життєвого циклу інформаційної системи; отримання практичних навичок використання і адаптації деяких найбільш відомих систем та бібліотек Data Mining.

знати:

- основні поняття та визначення інтелектуального аналізу даних;
- моделі та методи побудови моделей та аналізу залежностей у даних;
- сучасні програмні засоби для проектування і розробки систем інтелектуального аналізу даних;
- критерії порівняння моделей і методів інтелектуального аналізу даних.

вміти:

- обґрунтовувати й аналізувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при вирішенні практичних задач;
- використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних;
- створювати програми для інтелектуального аналізу даних при розв'язку конкретних практичних задач;
- аналізувати результати побудови та використання систем інтелектуального аналізу даних при вирішенні прикладних задач.

Інформаційний обсяг навчальної дисципліни

Тема № 1. Традиційні та сучасні технології інтелектуальної обробки даних.
Лекція №1. Основні поняття інтелектуальної обробки даних (ІОД). Сутність аналітичних технологій, поняття ІОД. Етапи та методи пошуку нових знань. Основні моделі інтелектуальних обчислень. Засоби програмної підтримки ІОД.

Лекція №2. Інтелектуальний аналіз даних – Data Mining. Класифікація задач Data Mining. Моделі Data Mining. Процес виявлення знань. Основні етапи аналізу. Підготовка початкових даних.

Лекція №3. Сховища даних та OLAP-технології. Концепція, технології побудови сховищ даних. Багатовимірна модель даних OLAP-системи.

Лекція №4. Стандарти Data Mining: CWM, CRISP, PMML, SQM/MM, OLE DB, DMAPI.

Лекція №5. Бібліотека Data Mining Xelopes. Архітектура бібліотеки. Класи моделі для Xelopes. Ієрархічні атрибути. Приклади використання бібліотеки Xelopes.

Тема № 2. Пошук асоціативних правил - Rules Mining та кластерний аналіз.

Лекція №6. Основні поняття Rules Mining. Програмні засоби пошуку асоціативних правил. Практичний аспект застосування технології асоціативних правил. Секвенціальний аналіз.

Лекція №7. Формальна постановка задачі кластеризації. Представлення результатів. Базові алгоритми кластеризації. Кластеризація по Гюстафсону-Кесселю.

Тема № 3. Візуальний аналіз даних - Visual Mining.

Лекція №8. Виконання візуального аналізу даних. Характеристики засобів візуалізації даних. Методи візуалізації. Методи геометричних перетворень. Відображення ікон. Методи, орієнтовані на піксели..

Тема № 4. Аналіз текстової інформації – Text Mining.

Лекція №9. Задача та етапи аналізу текстів Text Mining. Загальний опис процесу отримання понять з тексту. Методи класифікації текстових документів.

Лекція №10. Методи кластеризації текстових документів: ієрархічні, бінарні. Представлення текстових документів. Задача анотування текстів. Засоби аналізу текстової інформації.

Тема № 5. Особливості і ефективність генетичних алгоритмів

Лекція №11. Еволюційні технології та генетичні алгоритми. Концептуальні засади еволюційної теорії. Основні положення теорії генетичних алгоритмів.

Лекція №12. Моделі генетичних алгоритмів. Програмне забезпечення та сфери застосування генетичних алгоритмів. Мурашині алгоритми та генетичне програмування.

Тема № 6. Нечітка логіка.

Лекція №13. Нечіткі методи інтелектуального аналізу даних. Концепція нечітких обчислень. Нечітка логіка в системах Data Mining..

Лекція №14. Програмне забезпечення нечітких методів. Сучасна практика застосування нечітких методів в системах Data Mining.

Тема № 7. Розподілений аналіз даних.

Лекція №15. Системи мобільних агентів. Основні поняття, стандарти багатоагентних систем. Використання мобільних агентів для аналізу даних, агенти-аналітики, варіанти аналізу розподілених даних.

Лекція №16. Система аналізу розподілених даних (загальний підхід до реалізації системи, агент для збору інформації про базу даних, агент для збору статистичної інформації, агент для вирішення однієї задачі інтелектуального аналізу даних, агент для вирішення інтегрованої задачі інтелектуального аналізу даних)

Основна література:

Базова

1. Васютин С., Базы данных. Интеллектуальная обработка информации / С. Васютин, А. Гареев, Владимир Райх, Виктор Корнеев - М.:Нолидж,2000
2. Сегаран Т., Программируем коллективный разум / Т. Сегаран - СПб. : Символ-Плюс, 2008
3. Бакаев А. А., Методы организации и обработки баз знаний / А. А. Бакаев, В. И. Гриценко, Д. Н. Козлов - Киев : Наук. думка, 1993
4. Пайне А., Интеллектуальная обработка данных и хранилищ данных в SQL Server 2005 / А. Пайне - М. ; СПб. : Питер,2005

Допоміжна література

1. Han J., Data mining: Concepts and Techniques / Han J., Kamber M. - Morgan, Kaufmann Publishers. – 2001
2. Mitra S., Data Mining. Multimedia, Soft Computing, and Bioinformatics / Mitra S., Acharya T. - Hoboken, New Jersey. – 2003.