

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
 ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ФРАНКА
 ФАКУЛЬТЕТ ПРИКЛАДНОЇ МАТЕМАТИКИ ТА ІНФОРМАТИКИ
 КАФЕДРА ТЕОРІЇ ОПТИМАЛЬНИХ ПРОЦЕСІВ

**ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ (НА ВИБІР)
 ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ**

галузі знань 0403 – системні науки та кібернетика
 напрямку підготовки 6.040303 – системний аналіз
 факультету прикладної математики та інформатики

Виписка з навчального плану

Форма навчання	Курс	Семестр	Загальний обсяг (год.)	Всього аудит. (год.)	у тому числі (год.):			Самостійна робота (год.)	Контрольні (модульні) роботи (шт.)	Програмні проекти (шт.)	Курсові проекти (роботи), (сем.)	Залік (сем.)	Екзамен (сем.)
					Лекції	Лабораторні	Практичні						
Денна	4	8	72	42	14	28		30		1		8	-

АНОТАЦІЯ

Інтелектуальний аналіз даних (Data Mining, Knowledge Discovery in Data) - це процес аналітичного дослідження великих масивів інформації з метою виявлення нових реальних неочевидних нетривіальних закономірностей і систематичних взаємозв'язків між змінними, які згодом можна буде застосувати до нових наборів даних. Цей процес включає три основних етапи : дослідження, побудову моделі та її перевірку.

Data Mining - мультидисциплінарна область, яка виникла і розвивається на базі таких наук як прикладна статистика, розпізнавання образів, штучний інтелект, теорія баз даних та ін.

ЗМІСТ ПРОГРАМИ

Що таке інтелектуальний аналіз даних (Data Mining, Knowledge Discovery in Data). Дані. Інформація. Знання. . Історичний огляд. Виникнення, перспективи, проблеми Data mining. Поняття даних. Об'єкт і атрибут, вибірка, залежна і незалежна змінна. Типи шкал. Типи наборів даних. Поняття бази даних и СУБД. Поняття "інформація" і "знання", співставлення и порівняння цих понять.

Методи і стадії Data Mining. Стадії Data Mining і дії, що виконуються в рамках цих стадій. Класифікації методів Data Mining. Порівняльна характеристика деяких методів.

Задачі Data Mining. Задачі класифікації і кластеризації. Опис суті задач, процес розв'язування, методи розв'язування, застосування. Порівняння цих двох задач. Задачі прогнозування. Поняття часового ряду, його компонент, параметрів прогнозування, види прогнозів. Задача візуалізації даних.

Сфери застосування Data Mining. Основні сфери діяльності людини, де можна успішно застосовувати технології Data Mining. Поняття Web Mining, Text Mining, Call Mining.

Методи класифікації і прогнозування. Метод побудови дерева розв'язків для розв'язування задачі класифікації. Алгоритм „Кора” і C4.5. Основні ідеї методу опорних векторів і методу "найближчого сусіда". Переваги і недоліки цих методів. Приклади розв'язування задач.

Методи кластерного аналізу. Основи кластерного аналізу, математичні характеристики кластера. Дві групи ієрархічного кластерного аналізу: агломеративні і дивізімні методи.

Методи пошуку асоціативних правил. Суть задачі пошуку асоціативних правил. Алгоритм Аргіогі.

Процес Data Mining. Процес підготовки даних, поняття якості даних, засмічених даних, етапи очистки даних, редагування даних. Етапи процесу Data Mining, пов'язані з побудовою, перевіркою, оцінкою, вибором и корекцією моделей. Поняття "модель" і "моделювання".

Робота з пакетами See 5.0 та WizWhy. Технічний опис систем. Елементи інтерфейса користувача. Графічні можливості системи. Створення звітів. Налаштування параметрів каналу виводу. Виконання лабораторних робіт.

ЛІТЕРАТУРА

1. **А.А.Барсегян.** Технології аналізу даних: Data Mining, Visual Mining. Text Mining, OLAP.// **А.А.Барсегян, М.С.Куприянов, В.В.Степаненко І.І.Холод** Санкт-Петербург. „БХВ-Петербург”. 2007.-376с.
2. **В.Дюк, А.Самойленко.** Data Mining: Ученый курс. // **В.Дюк, А.Самойленко –** СПб: Питер. 2001. - 368с.
3. **В.П.Боровиков.** STATISTICA.Искусство анализа данных на компьютере: Для профессионалов – СПб.: Питер, 2003 – 688 с.
4. **Jiawei Han.** Data Mining: Concepts and Tecniques.// **Jiawei Han, Micheline Kamber.** Morgan Kaufmann Publishers, 2001.
5. **Ian H.Witten.** Data Mining. // **Ian H.Witten, Eibe Frank.** Academic Press, 2000.

Програму склала ст.викл.
кафедри теорії оптимальних процесів

О.Гнатишин