

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені В.Н. КАРАЗІНА

Кафедра міжнародних економічних відносин імені Артура Голікова

КОМПЛЕКС НАВЧАЛЬНО-МЕТОДИЧНОГО ЗАБЕЗПЕЧЕННЯ

з дисципліни
«ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ В СИСТЕМАХ
ЕЛЕКТРОННОЇ КОМЕРЦІЇ»

рівень вищої освіти перший (бакалаврський)
галузь знань 29 «Міжнародні відносини»
спеціальність 292 «Міжнародні економічні відносини»
освітня програма «Міжнародна електронна комерція»,
вид дисципліни обов'язкова
факультет міжнародних економічних відносин та туристичного бізнесу

Укладач: ст. викл. Ревякін Г.В.

Розширений план лекцій

Тема 1. Вступ до інтелектуального аналізу даних

1. Визначення Data Mining і область застосування.
2. Задачі, моделі та методи Data Mining.
3. Поняття Business Intelligence.

Тема 2. Предмет, метод і основні поняття статистики

1. Виникнення, формування та розвиток статистики.
2. Предмет статистики.
3. Теоретичні основи статистики.
4. Зв'язок статистики з іншими суспільними науками.
5. Основні категорії статистики: статистична сукупність, статистична закономірність, закон великих чисел.
6. Статистичні ознаки та їх класифікація, варіація ознак.
7. Статистичні закономірності та форми їх вияву. Загальне поняття про статистичну методологію.
8. Основні розділи статистичної науки.

Тема 3. Етапи роботи з великими об'ємами даних

1. Цикл одержання, попередньої обробки, аналізу даних, інтерпретації результатів та їхнього використання.
2. Етапи процесу Data Mining, пов'язані з побудовою, перевіркою, оцінкою, вибором и корекцією моделей.
3. Методи первісної обробки даних.
4. Інструментальні засоби Data Mining.

Тема 4. Випадкові величини

1. Основні розділи аналізу даних.
2. Класифікація змінних.
3. Групування даних.
4. Головні задачі попередньої обробки даних.
5. Квантили та відсоткові точки.

Тема 5. Числові характеристики розподілу

1. Поняття ряду розподілу та його елементи.
3. Форми рядів розподілу: симетричні та асиметричні ряди. Основні характеристики ряду розподілу: середня, мода, медіана, квартилі, децилі.
4. Суть варіації, необхідність її статистичного вивчення.
5. Розмах варіації, середнє лінійне відхилення, середній квадрат відхилення (дисперсія) та середнє квадратичне відхилення, коефіцієнт варіації.

Тема 6. Перевірка статистичних гіпотез

1. Поняття про статистичну гіпотезу.
2. Нульова та альтернативна гіпотези.

3. Ризики при перевірці гіпотез.
4. Рівень істотності. Поняття статистичного критерію. Параметричні та непараметричні критерії.
5. Алгоритм перевірки гіпотез.

Тема 7. Коваріаційний та кореляційний аналіз

1. Види взаємозв'язків між явищами.
2. Функціональний та кореляційний зв'язки.
3. Дисперсійний аналіз.
4. Поняття про кореляційний-регресійний аналіз, основні завдання кореляційного аналізу. Передумови кореляційного аналізу.
5. Види кореляційних залежностей за формою та напрямом зв'язку. Аналіз простої лінійної кореляції.
6. Коефіцієнт детермінації, аналіз простої нелінійної кореляції.

Тема 8. Модель парної регресії: властивості коефіцієнтів та перевірка гіпотез

1. Аналіз множинної регресії.
2. Побудова рівняння регресії.
3. Методика обчислення параметрів рівняння регресії.
4. Парні, часткові та множинний коефіцієнт регресії.
5. Часткові та множинний коефіцієнти детермінації.

Тема 9. Модель множинної регресії

1. Множинна регресія.
2. Специфікація багатofакторної моделі.
3. Мультиколінеарність та методи її усунення.
4. Помилки специфікації множинної регресії та практичні наслідки мультиколінеарності.
5. Оцінка якості моделі множинної регресії.

Тема 10. Гетероскедастичність та автокореленість випадкового члена

1. Поняття гомоскедастичності та гетероскедастичності залишків.
2. Наслідки порушень припущення про гомоскедастичність.
3. Методи виявлення гетероскедастичності.
4. Трансформування початкової моделі з гетероскедастичністю.
5. Зважений метод найменших квадратів.
6. Оцінювання параметрів регресії за допомогою узагальненого методу найменших квадратів.