

Розділ 1. Хімічна інформація. Тема 1. Джерела хімічної інформації Джерела хімічної інформації, бази даних і пакети прикладних програм (Science Citation Index, CSD тощо). Хімічні періодичні видання. Друковані та електронні версії. Імпакт-фактори журналів. Пошук інформації в Інтернет. Статистика, хімічна інформатика, хемометрія Експериментальне визначення властивостей речовин. Джерела довідкової інформації. Розрахунки властивостей речовин. Емпіричні та неемпіричні розрахунки. Тема 2. Властивості речовин. Бази даних властивостей. Тема 3. Кореляції структура-властивість. QSPR Побудова моделей залежності структура-властивість. Прогнозування властивостей на основі дескрипторних моделей. Тема 4. Основні поняття теорії моделювання. Загальні відомості про моделювання. Сутність і стадії математичного моделювання. Моделі-технологічних процесів. Ієрархічна структура хімічного виробництва. Основні положення системного аналізу. Тема 5. Ризик-орієнтований аналіз. Ризик. Аналіз ризиків хіміко-технологічних процесів. Місце ризиків у системі управління якістю. Контрольні точки. Коригувальні дії. Принцип документованої інформації. Тема 6. Прикладне програмне забезпечення. Пакет ACD Labs Основний функціонал. Редактор формул chemsketch. Пакети обробки аналітичних даних. Тема 7. Локальні бази даних. Загальна характеристика локальних баз даних. Пакет ISIS Base. Інше програмне забезпечення. Представлення структурної інформації у базах даних. Тема 8. Бази даних Reaxys, Scifinder. Комерційні бази даних хімічної інформації. Бази даних з вільним доступом. Тема 9. Некомерційне програмне забезпечення Пакети програм доступні безкоштовно. Безкоштовні версії програмного забезпечення з обмеженим функціоналом, версії для навчання та ознайомлення.

Розділ 2. Статистичні методи в хімії Тема 10. Похибки у експериментальних даних Види похибок. Емпіричні та теоретичні розподіли. Закон розповсюдження похибок. Випадкові та системні похибки. Методи перевірки. Тема 11. Представлення даних. Первинне представлення даних Дескриптивне представлення даних. Гістограми. Результат вимірювань як випадкова величина Генеральна сукупність та вибірка. Вибіркові оцінки. Моменти. Середнє. Стандартне відхилення. Дисперсія. Коваріаційні матриці. Коефіцієнти кореляції. Перетворення даних (масштабування, автомасштабне перетворення). Тема 12. Статистична обробка даних Задачі обробки первинних експериментальних даних: дослідження однорідності вибірки, визначення функції розподілу, кореляцій між змінними, класифікація, факторний аналіз. Структурна і параметрична ідентифікація моделей, перевірка адекватності. Статистичні розподіли випадкової величини. Дискретні й неперервні випадкові величини. Біноміальний розподіл. Розподіли неперервних величин: рівномірний, Гаусса, Лапласа, 2. Центральна гранична теорема. Метод максимуму правдоподібності. χ^2 Пуассона, Функція правдоподібності. Правдоподібні оцінки параметрів генеральної сукупності при нормальному та Лапласівському розподілах похибок. Тема 13. Перевірка статистичних гіпотез Задача перевірки статистичних гіпотез. Схема перевірки гіпотези. Помилки I та II 2, χ родів. Потужність критеріїв. Перевірка гіпотез про функції розподілу. Критерій графічні способи перевірки гіпотез про функції розподілу Тема 14. Основи кореляційного та регресійного аналізу Кореляційний аналіз. Приклади кореляцій в хімії, значення кореляцій. Принцип лінійності вільних енергій як основа багатьох хімічних кореляцій. Теоретичні засади методу найменших квадратів (МНК) та статистичні властивості оцінок МНК. Розрахункова схема МНК. Вибір найкращого набору регресорів: методи всіх регресій, покрокової регресії, вилучення регресорів. Приклади використання МНК у хімічних задачах. Лінійний та нелінійний МНК як приклад некоректної задачі (теоретичний аналіз та приклади), мультиколінеарність. Її формальні та неформальні причини. Способи подолання мультиколінеарності: α -регуляризація за Тихоновим, застосування ортогональних поліномів, сингулярний розклад. Типові приклади математично некоректних задач в хімії. Вплив викидів на оцінки МНК. Уявлення про робастні оцінки. Тема 15.

Класифікація та кластерний аналіз Види класифікацій і їх значення для аналізу даних. Типи ознак. Міри схожості об'єктів. Класифікація без навчання. Ієрархічна класифікація, дендрограми. Найпростіші алгоритми ієрархічного кластерного аналізу сукупності об'єктів. Проблема стійкості класифікації. Факторний аналіз. Характеристика моделей з латентними змінними. Кореляційна та коваріаційна матриці – об'єкт факторного аналізу. Формулювання задачі факторного аналізу. Матриця навантажень, вектори характерностей і загальностей. Основна факторна теорема. Експлораторний та конфірматорний факторний аналіз. Алгоритми факторного аналізу. Факторний аналіз хроматографічних даних.