

КОМПЛЕКС ЛАБОРАТОРНИХ РОБІТ «ПОБУДОВА ПРОГНОСТИЧНИХ МОДЕЛЕЙ ДЛЯ ОПИСУ рKa ЛІКАРСЬКИХ СПОЛУК»

Тема: «Напівемпіричні методи розрахунку геометрій молекул, Розрахунок молекулярних дескрипторів»

- оптимізуйте геометрію молекул (методи: AM1 або PM3);
- Знайдіть атомні заряди на гетеро атомах;
- Обчисліть коефф. кореляції між рKa і зарядами на гетероатомах базової структури;
- Побудуйте графік залежності рKa від заряду, який найсильніше корелює з рKa;
- Випадковим чином поділіть первинну вибірку на «обучаючу» (80% молекул) та «тестову» (решта);
- Для навчаючої вибірки, за допомогою програми PADEL (або DRAGON) розрахуйте набір дескрипторів.

- 1) У звіті опишіть геометричну структуру сполук навчаючої вибірки (номер молекули в вихідній таблиці і рисунок по результатам розрахунку методом AM1).
- 2) обчисліть дескриптори

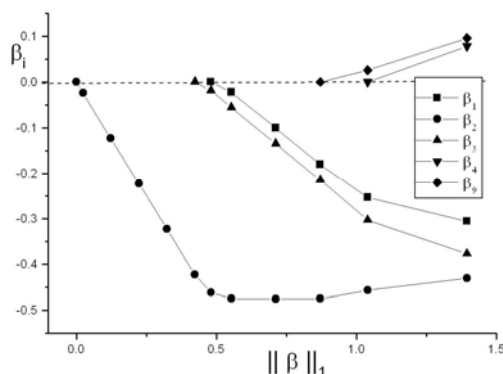
Тема: «Регресійні моделі QSAR»

Метод найменших квадратів

- 3) Знайдіть і опишіть найбільш значущі (корелюючі із рK) дескриптори.
- 4) побудуйте регресійні рівняння, що включають 1 і 2 дескриптора, розрахуйте величини R^2 , Q^2 і стандартні відхилення.
- 5) для кожної регресії побудуйте графіки залежності «теорія (розрахунок) – експеримент» та «теорія (передбачення) – експеримент».

Метод LASSO

- 6) Реалізуйте розрахунок з різними величинами параметрів L_1 -регуляризації. Підберіть параметр регуляції так, щоб отримані рівняння залежали від одного та двох параметрів. Чи ідентичні вони тем, що були отримані в п. 3 і 4?
- 7) побудуйте графік залежності від параметрів регресії від або $|\beta|$. Він може мати наприклад такий вигляд:



8) Опишіть якість отриманих регресійних рівнянь.

Метод регресії на головних компонентах (Principal Component Regression, PCR)

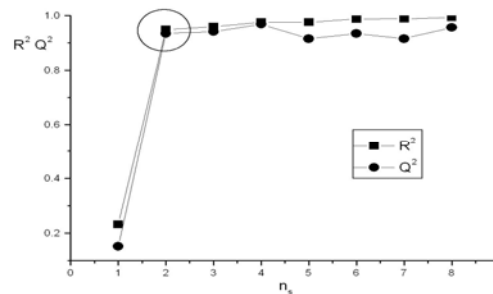
9) Реалізуйте PCR з $n_s = 1, 2, 3, 4, 5, 6$ сингулярними числами.

10) Для кожного з них побудуйте графік «теорія (розрахунок) – експеримент»

11) Для кожного з них побудуйте графік «теорія (передбачення) – експеримент».

12) запишіть відповідні коефіцієнти детермінації (R^2 , Q^2) та стандартні відхилення.

13) Побудуйте результуючий графік (R^2 , Q^2 , як функція n_s) та надайте відповідну інтерпретацію. Цей графік може мати, наприклад, такий вид



Тема: «Факторний аналіз»

14) Проведіть факторний аналіз та опишіть структуру двох найбільш значимих факторів. Якщо за результатами факторного аналізу можна запропонувати ще будь які регресійні рівняння (відмінні від тих, що були отримані раніше) то побудуйте та дослідіть їх.

Тема: «Метод штучних нейронних мереж»

15) На основі попередньої роботи виберіть 4 найбільш значущих дескриптора;

16) Налаштуйте мережу з 4 нейронів. Реалізуйте три варіанти мереж: 2 дескриптора, 3 дескриптора, 4 дескриптора;

17) Обчисліть (для цих трьох варіантів) передбачувані значення властивостей (Predict)

18) Збережіть передбачені значення властивостей (**Output** → **Save**). Збережіть файл.

19) Імпортуйте ці дані в EXCEL та обчисліть величини кореляції та стандартних відтворень.

20) Збережіть найкращу налаштовану мережу у html форматі (**export** → **web page** → **save**)

21) Протестуйте нейронну мережу за допомогою тестової вибірки. Підрахуйте кількість коректних прогнозів.