STATYSTYKA I CHEMOMETRIA W ANALITYCE CHEMICZNEJ

Instrukcja do ćwiczeń laboratoryjnych

Ćwiczenie 3

**1. Cel ćwiczenia:**

Celem ćwiczenia jest nabycie umiejętności poprawnego wyznaczenia równania regresji wielokrotnej do modelowania zależności liniowych dla więcej niż jednej zmiennej niezależnej, oraz poprawne zinterpretowanie uzyskanych wyników a także ocena zdolności prognostycznej uzyskanego modelu.

**2. Zagadnienia do samodzielnego opracowania:**

Pojęcie modelu matematycznego. Zasada działania metody regresji wielokrotnej. Istotność statystyczna modelu. Jakość dopasowania modelu. Zdolność prognostyczna modelu. Rachunek macierzowy (mnożenie macierzy, transponowanie macierzy, odwracanie macierzy).

**3. Przebieg ćwiczenia:**

*Opis zadania*

W wyniku dodatkowych badań przeprowadzonych na próbkach przez oddzielne laboratorium, zidentyfikowano kolejną substancję zanieczyszczającą narkotyk. Był to związek organiczny z grupy dioksyn, który stanowi silną truciznę. Wykrycie jego obecności w narkotyku stanowi zatem wysoki priorytet. Niestety eksperyment pozwalający na zbadanie ilości nowego związku jest niezwykle czasochłonny, przez co szybkie przebadanie kolejnych próbek jest niemożliwe. Twoim zadaniem jest sprawdzenie, czy jego obecność związana jest silnie z obecnością któregoś z pozostałych zanieczyszczeń. W celu zidentyfikowania potencjalnie powiązanego z dioksyną zanieczyszczenia należy obliczyć odpowiednią macierz korelacji oraz poprzeć swój wybór dodatkową analizą HCA. Należy także sprawdzić czy z obecnością dioksyny nie jest związana żadna z otrzymanych w wyniku analizy PCA głównych składowych. Dla wybranego zanieczyszczenia oraz wybranej głównej składowej należy wykonać modele regresji, opisujące ilościowo zawartość dioksyny w narkotyku. Wyniki należy porównać i przedstawić wraz z interpretacją.

Dla ułatwienia identyfikacji poszczególnych zanieczyszczeń tabelę z poprzednich zadań zamieszczono poniżej.

Tabela 1. Nazwy substancji, których obecność w próbkach amfetaminy badano w laboratorium.

|  |  |
| --- | --- |
| Symbol | Nazwa |
| X1 | 1,2-diphenyletylamine |
| X2 | 4-benzylpyrimidine |
| X3 | Alfa-methyldiphenetylamine |
| X4 | N,N-dibenzyloamine |
| X5 | 2,6-dimethyl-3,5-diphenylpyridine |
| X6 | 4-methyl-5-phenylpyrimidine |
| X7 | 1,3-dimethyl-2-phenylnaphthalene |
| X8 | 1,3-diphenyl-2-propylamine |
| X9 | N-benzoylamphetamine |
| X10 | 2,6-diphenyl-3,4-dimethylpyridine |

Zadanie:

Na zajęciach zostanie przedstawiony schemat, zawierający informacje jakie kroki należy podjąć aby rozwiązać problem przedstawiony w treści zadania. Studenci mają za zadanie przeprowadzenie analizy poprzez:

* przygotowanie odpowiednich algorytmów analiz w programie KNIME (przygotowane algorytmy zatwierdza prowadzący),
* import danych do programu KNIME,
* odpowiednią konfigurację niezbędnych do dalszej pracy NODÓW oraz wykonanie wszystkich kroków analizy,
* zapisanie oraz poprawną interpretację wyników.

**4. Sprawozdanie:**

Wyniki uzyskane na zajęciach wraz z interpretacją.

**5. Literatura**

A. Łomnicki, „Wprowadzenie do statystyki dla przyrodników”, Wydanie trzecie uzupełnione, Wyd. Naukowe PWN, Warszawa, 2005

J. Mazerski, „Podstawy chemometrii”, Wydawnictwo Politechniki Gdańskiej, Gdańsk, 2000