**Opis problemu:**

Problem polega na diagnozowaniu raka piersi u pacjentów na podstawie 9 atrybutów. Na wyjściu jest informacja o tym czy w poszczególnych rekordach ( u poszczególnych pacjentów) choroba raka piersi jest łagodna czy złośliwa. Więcej informacji poniżej oraz w pliku *breast-cancer-description.txt*w katalogu projektu.

**Źródło:**

Creators:   
1. Dr. William H. Wolberg, General Surgery Dept.   
University of Wisconsin, Clinical Sciences Center   
Madison, WI 53792   
wolberg **'@'** eagle.surgery.wisc.edu

2. W. Nick Street, Computer Sciences Dept.   
University of Wisconsin, 1210 West Dayton St., Madison, WI 53706   
street **'@'** cs.wisc.edu 608-262-6619

3. Olvi L. Mangasarian, Computer Sciences Dept.   
University of Wisconsin, 1210 West Dayton St., Madison, WI 53706   
olvi **'@'** cs.wisc.edu   
  
Donor:Nick Street

**Informacje o zbiorze danych:**

Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image.   
  
Separating plane described above was obtained using Multisurface Method-Tree (MSM-T) [K. P. Bennett, "Decision Tree Construction Via Linear Programming." Proceedings of the 4th Midwest Artificial Intelligence and Cognitive Science Society, pp. 97-101, 1992], a classification method which uses linear programming to construct a decision tree. Relevant features were selected using an exhaustive search in the space of 1-4 features and 1-3 separating planes.   
  
The actual linear program used to obtain the separating plane in the 3-dimensional space is that described in: [K. P. Bennett and O. L. Mangasarian: "Robust Linear Programming Discrimination of Two Linearly Inseparable Sets", Optimization Methods and Software 1, 1992, 23-34].   
  
This database is also available through the UW CS ftp server:   
ftp ftp.cs.wisc.edu   
cd math-prog/cpo-dataset/machine-learn/WDBC/

**Informacje o atrybutach:**

1) ID number   
2) Diagnosis (M = malignant, B = benign)   
3-32)   
Ten real-valued features are computed for each cell nucleus:   
  
a) radius (mean of distances from center to points on the perimeter)   
b) texture (standard deviation of gray-scale values)   
c) perimeter   
d) area   
e) smoothness (local variation in radius lengths)   
f) compactness (perimeter^2 / area - 1.0)   
g) concavity (severity of concave portions of the contour)   
h) concave points (number of concave portions of the contour)   
i) symmetry   
j) fractal dimension ("coastline approximation" - 1)