Task-1 (easy)

Determining the type of pathology from tabular data

In this task you need to determine whether a breast tumor is malignant (M) or benign (B).

Features are calculated based on the digitized image of the tumor. They describe the characteristics of the cell nuclei present in the image.

Data are here

The target variable is Diagnosis.

The goal of the project is to build machine learning models to solve the classification problem and select the best model. Evaluate the performance of this model on a validation sample and identify the most significant features for classification. Present the results and defend your solution.

Task-2 (hard)

Segmentation of low-grade glioma using MRI images of the brain

This project proposes to segment low-grade gliomas based on a dataset from the cancer imaging archive The Cancer Genome Atlas (TCIA). The dataset contains MR images of the brain along with manually generated FLAIR anomaly segmentation masks.
Tumor genomic clusters and patient data are presented in the data.csv file.

The data is at the link:

https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation/download?datasetVersionNumber=2

Based on this dataset, a kaggle competition was created, in which you are invited to participate.

The goal of the project is to build a segmentation model to solve the problem of segmenting a tumor using an anomaly mask. Compare the obtained metrics with existing solutions and present the best model and your results