

# Лабораториска вежба 5 - Класификација на графови

## Податочно множество - DD

Податочното множество DD се состои од 1 178 протеински структури претставени како графови. Јазлите претставуваат аминокиселини. Два јазли се поврзани ако растојанието помеѓу соодветните аминокиселини е помало од 6 ангстроми (0.6 нанометри). Целта е да се одреди дали дадена протеинска структура е ензим или не. За вчитување на множеството искористете го следниот код:

```
from torch_geometric.datasets import TUDataset  
  
data = TUDataset(root='../data/TUDataset/DD', name='DD', use_node_attr=True)
```

## Задачи

### Задача 1 - Класификација на графови со GraphSAGE (25 поени)

Направете модел за класификација на графови кој го имплементира GraphSAGE. Првиот дел на моделот се 2 или 3 слоеви во GraphSAGE, по што се надоврзуваат 1, 2 или 3 скриени FC слоеви за класификација на графовите. Направете класификација на графовите преку споредба на различните големини и комбинации на слоеви на моделите. Поделете ги графовите на графови за тренирање, валидација и тестирање. Истренираниот модел евалуирајте го со метриките: точност (`accuracy_score`), прецизност (`precision_score`), одзив (`recall_score`) и F1-мерка (`f1_score`).

### Задача 2 - Класификација на графови со GCN (25 поени)

Направете и втор модел за класификација на графови кој го имплементира GCN. Првиот дел на моделот се 2 или 3 слоеви во GCN, по што се надоврзуваат 1, 2 или 3 скриени FC слоеви за класификација на графовите. Направете класификација на графовите преку споредба на различните големини и комбинации на слоеви на моделите. Користете ја истата поделба како во претходната задача. Направете евалуација на истиот начин како во претходната задача. Со кој метод се добиваат подобри резултати?