



Project MAP583: Graph classification by GNN

DONG Tian, GUO Shijia, GUO Yanzhu, NI Runbo

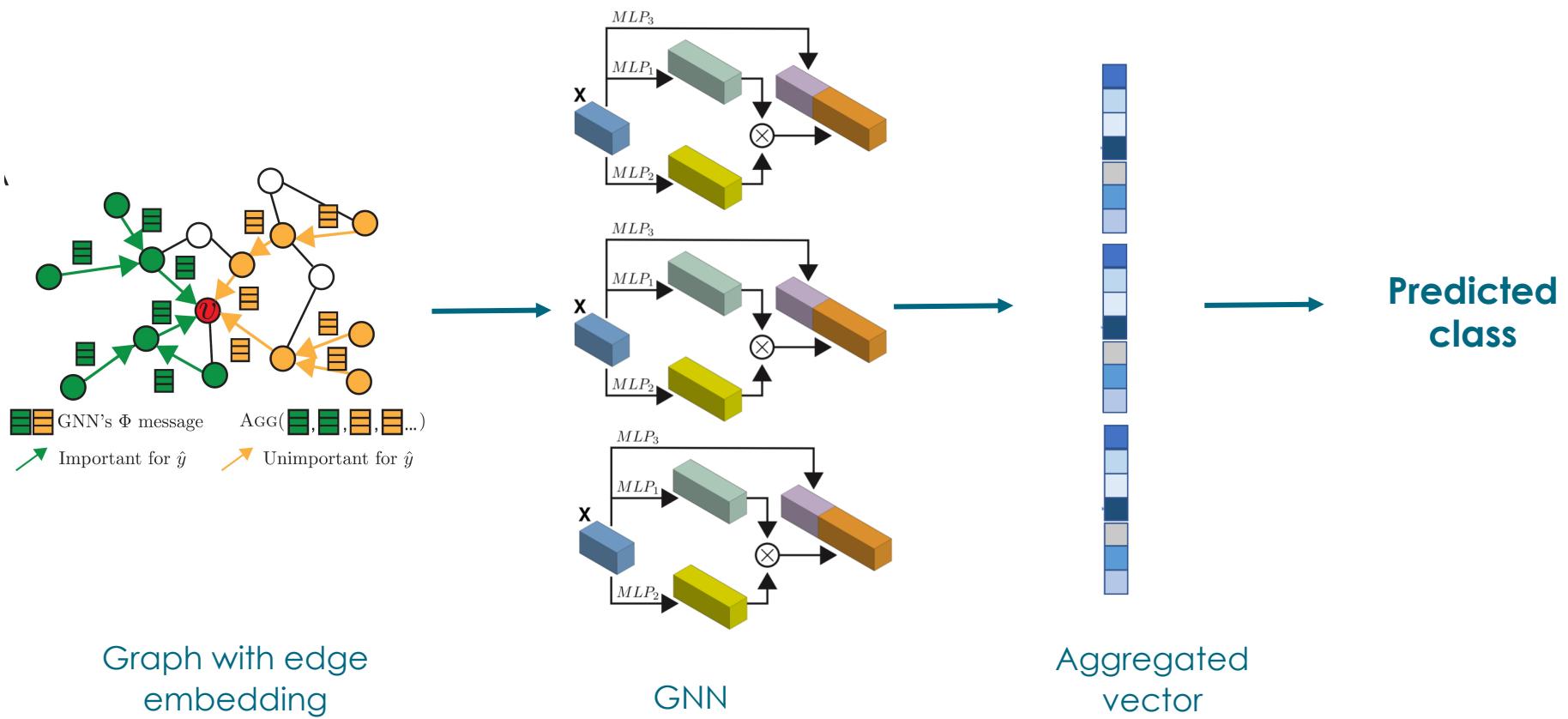
GitHub: https://github.com/pqros/MAP583_projet

Outline

- **Introduction: learning task**
- **Datasets**
- **Graph Neural Networks (GNN)**
- **Results**



Learning task

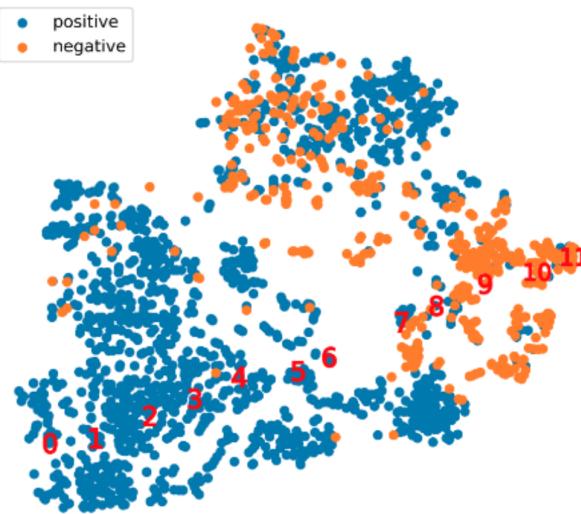


Datasets

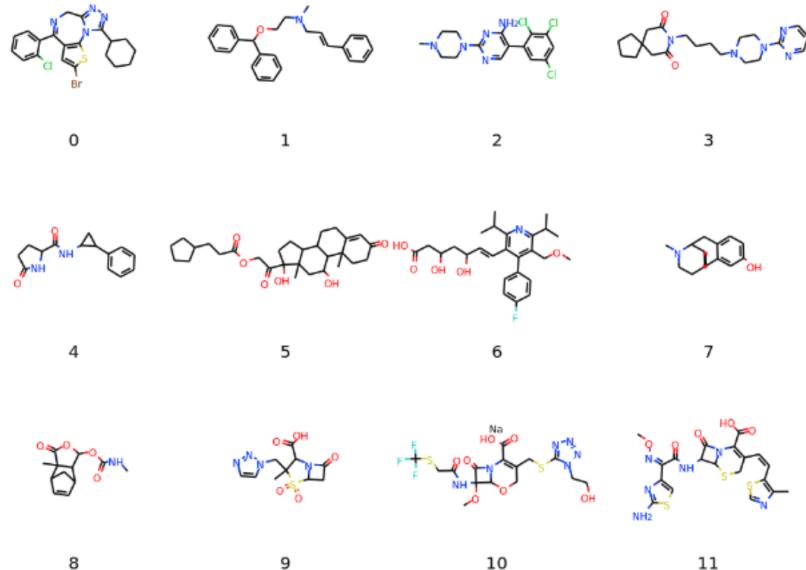


Dataset	Category	Size	# Tasks
BBBP	Physiology	2039	1
Tox21		7831	12
ToxCast		8575	617
SIDER		1427	27
ClinTox		1478	2
MUV	Biophysics	93087	17
HIV		41127	1
BACE		1513	1

Datasets

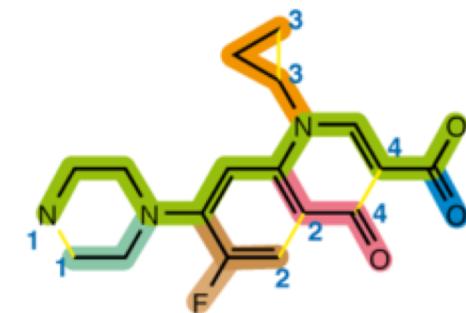


BBBP



Example: molecules
in BBBP

C

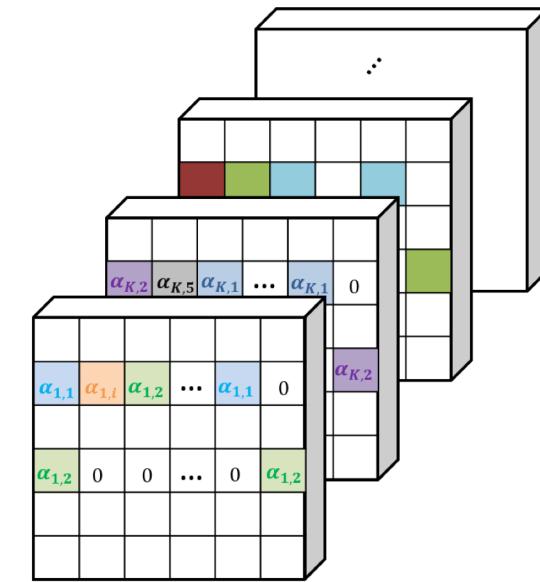
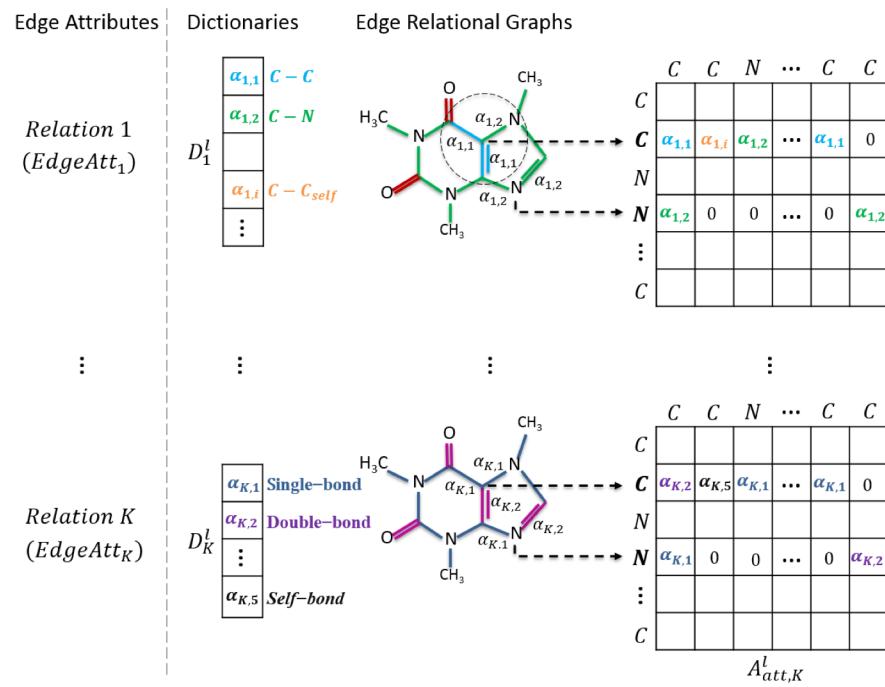
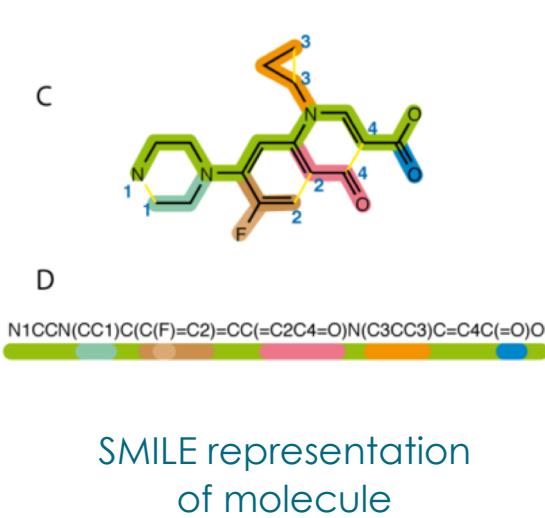


D

N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O

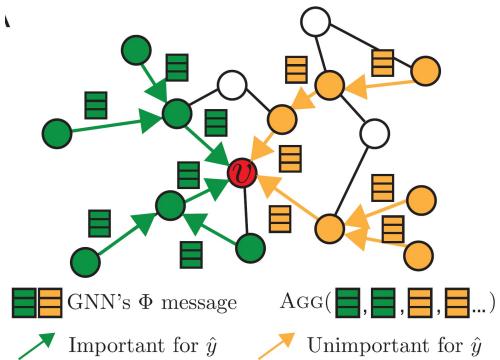
SMILE representation
of molecule

Datasets

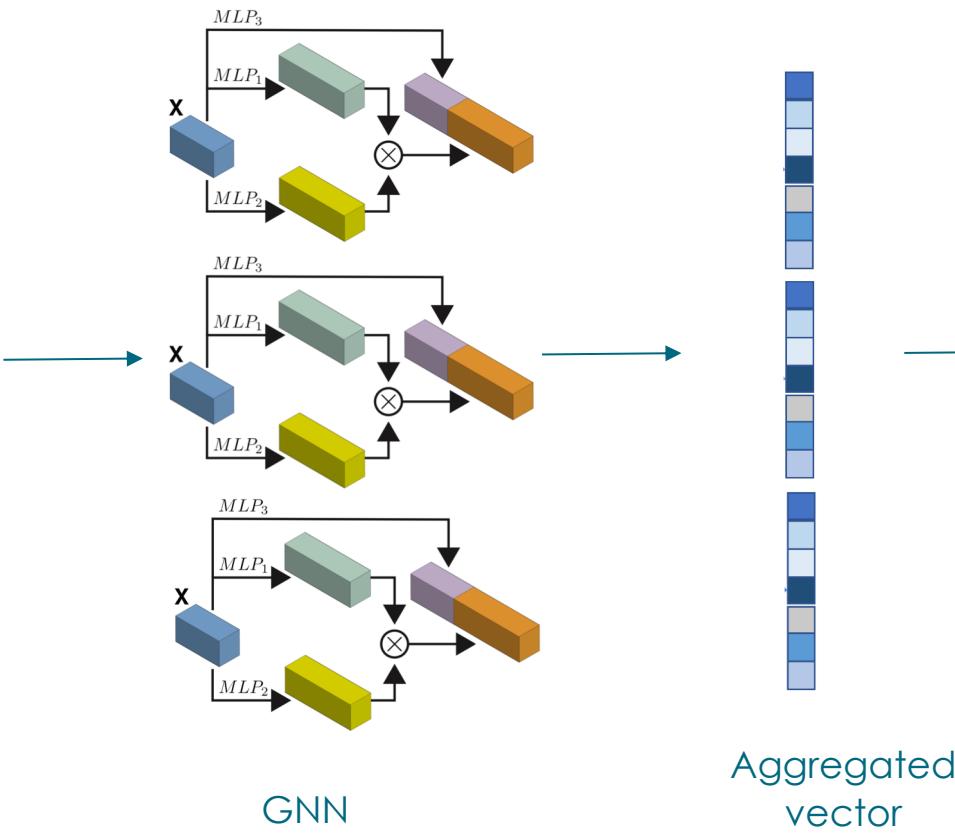


Multi-channel
Adjacency matrix with
edge embedding

GNN



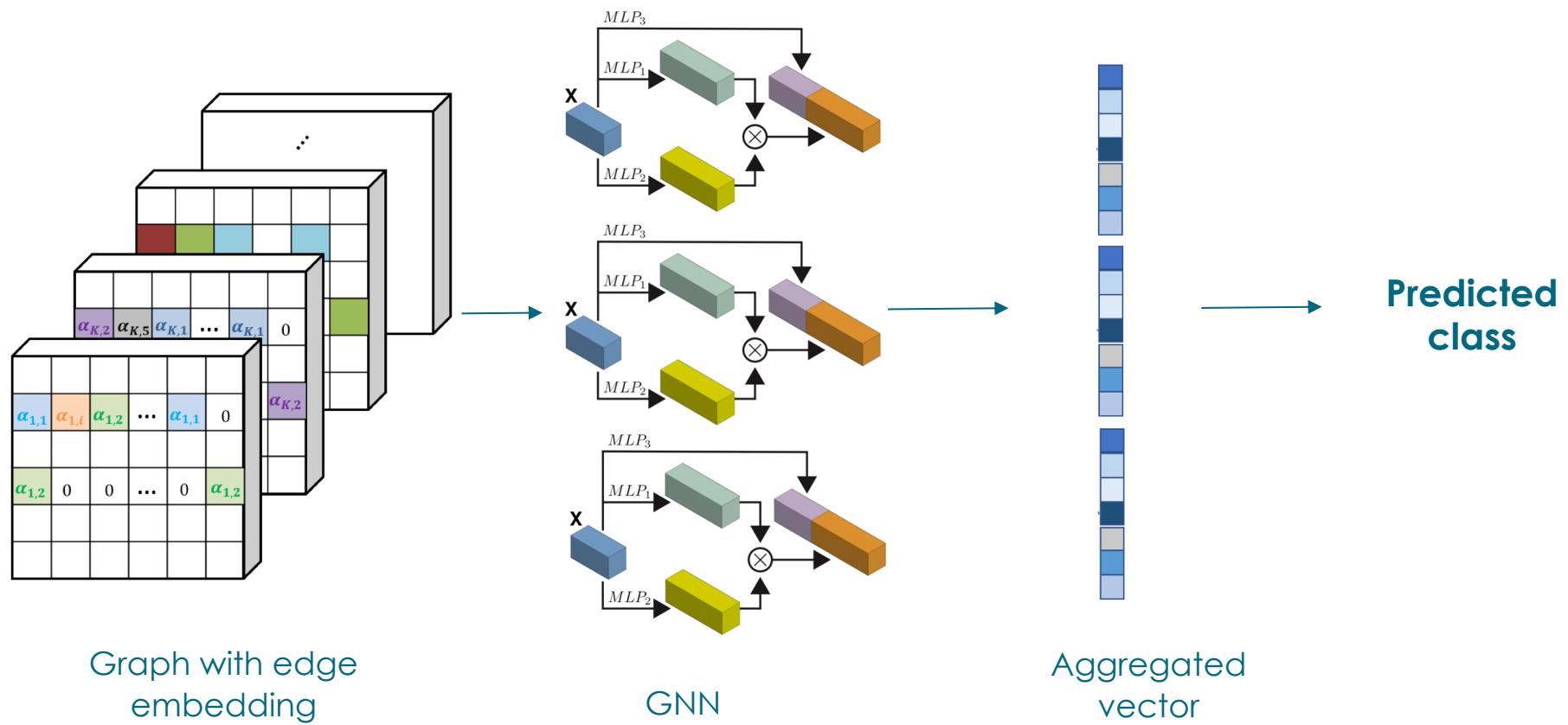
Graph with edge
embedding



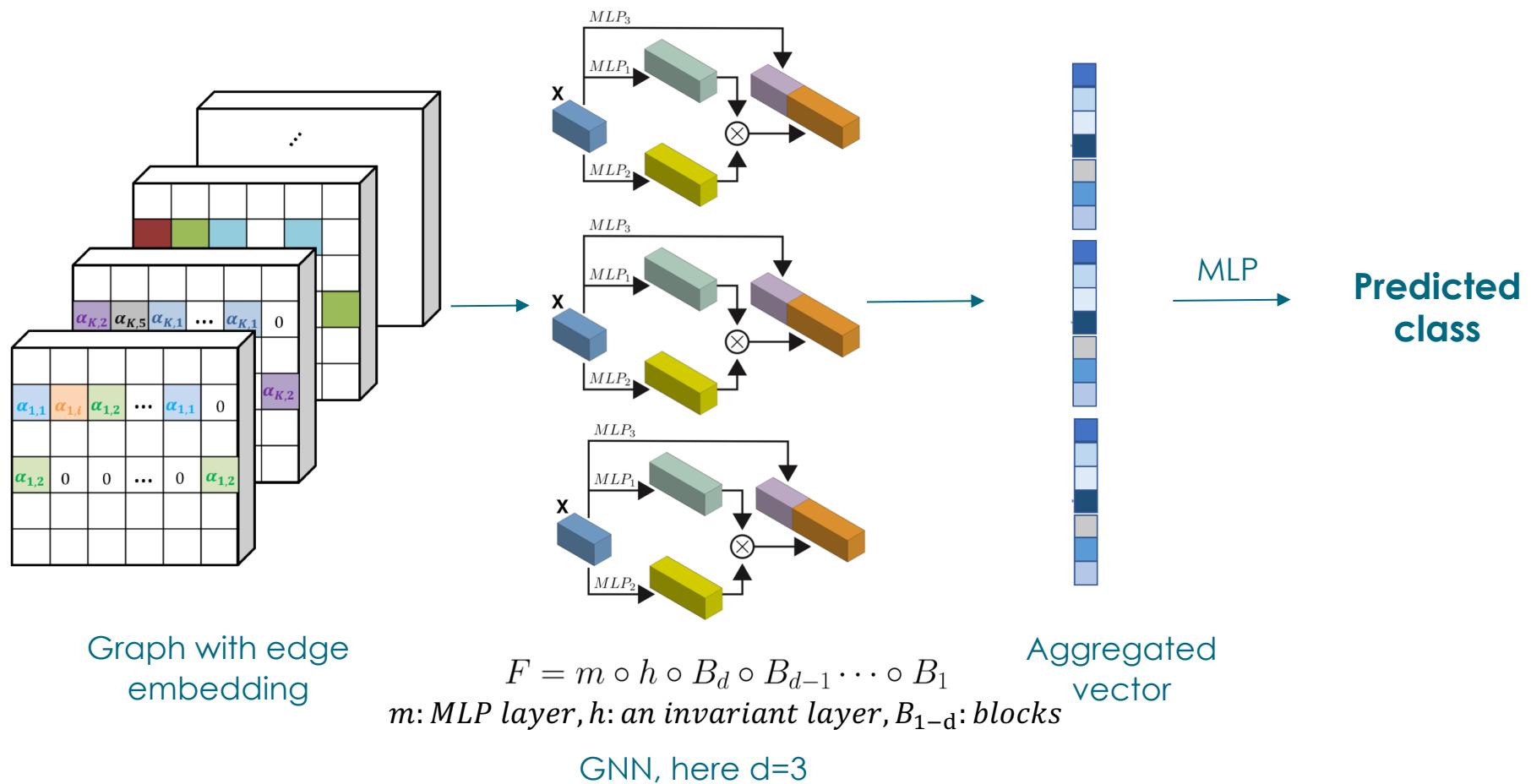
Aggregated
vector

Predicted
class

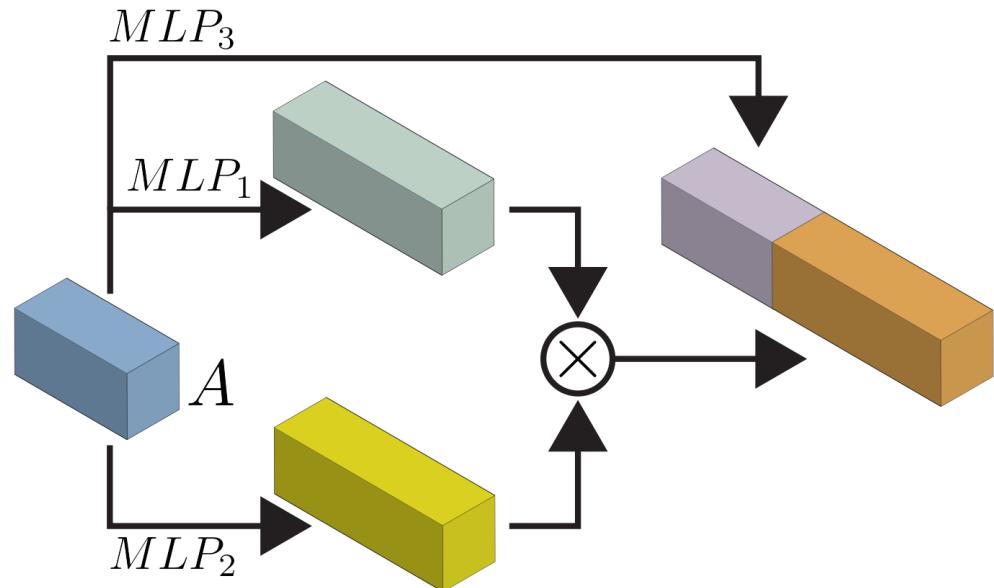
GNN



GNN



GNN



```
1 ▶ def forward(self, input):
2     mlp1 = self.MLP1(input)
3     mlp2 = self.MLP2(input)
4     mlp3 = self.MLP3(input)
5
6     mult = torch.matmul(mlp1, mlp2)
7
8     return torch.cat((mult, mlp3), dim=1)
```

A basic block B_i of the GNN.

The block consists of (i) applying 3 different MLPs to the feature dimension of the input (ii) multiplying the output of two MLPs feature-wise and (iii) concatenating the output of the last MLP.

Results



Dataset	Category	Size	# Tasks	Our model	SOTA
BBBP	Physiology	2039	1	0.69	0.68
Tox21		7831	12	0.70	0.78
ToxCast		8575	617	0.57	0.66
SIDER		1427	27	0.59	0.63
ClinTox		1478	2	0.70	0.72
MUV	Biophysics	93087	17	0.70	0.81
HIV		41127	1	0.79	0.80
BACE		1513	1	0.71	0.84

ROC-AUC (%) results

Overfitting !



Thanks for your attention!