b

UNIVERSITÄT
BERN

UNIVERSITÄT
BERN

Matching of Matching-Graphs

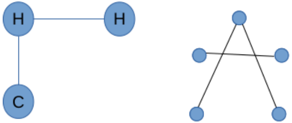
A Novel Approach for Graph Classification

Mathias Fuchs, Kaspar Riesen
10.12.2020

Graphs and Graph Matching

UNIVERSITÄT BERN

- A Graph consists of a set of nodes and a set of edges. Both of them can contain additional attributes.



- Graphs are a versatile alternative to feature vectors
- Various graph matching procedures have been proposed over the years
  - Graph Kernels
  - Spectral Methods
  - Graph Edit Distance
  - Graph Convolutional Neural Networks

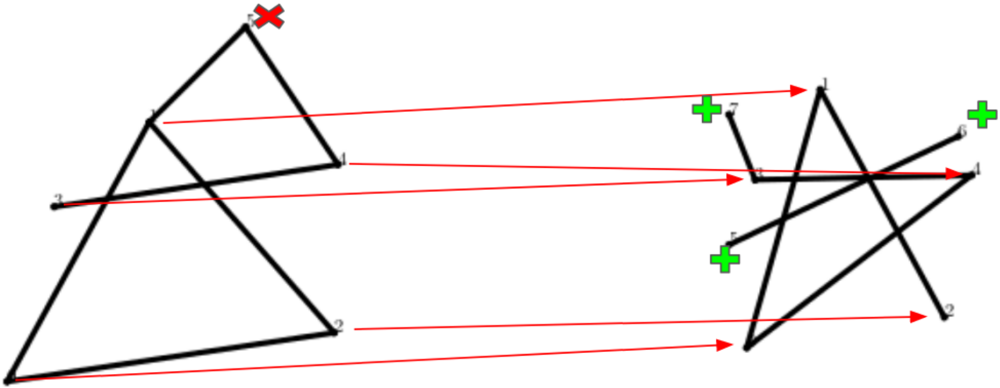
Graph Edit Distance

Definition

- Given two graphs g_1 and g_2, the basic idea of graph edit distance is to transform g_1 into g_2 using some *edit operations*
- Default set of edit operations
  - Insertion: A node is inserted. Denoted as $(\varepsilon \to v)$, where v is the inserted node.
  - Deletion: A node is deleted. Denoted as $(u \to \varepsilon)$, where u is the deleted node.
  - Substitution: A node is substituted with another node. Denoted as $(u \to v)$, where u is a node of g_1 and v a node of g_2
  - Similarly for edge edit operations

Graph Edit Distance

Example



The corresponding edit path is $= \{0 \rightarrow 0, 1 \rightarrow 1, 2 \rightarrow 2, 3 \rightarrow 3, 4 \rightarrow 4, 5 \rightarrow \varepsilon, \varepsilon \rightarrow 5, \varepsilon \rightarrow 6, \varepsilon \rightarrow 7\}$

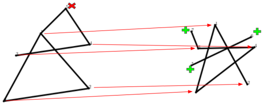
Matching-Graphs

Idea

- The general idea is to build a small set of graphs, that represent a certain class.
- This is done by formalizing the information of a given edit path between two graphs g_i and g_j in a new data structure, called the **matching-graph** $m_{g_i \times g_j}$.
- For each edit path $\lambda(g_i, g_j)$, two matching-graphs $m_{g_i \times g_j}$ and $m_{g_j \times g_i}$ are eventually built (for the source and the target graph g_i and g_j, respectively)

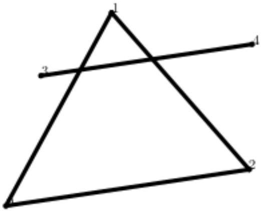
Matching-Graphs

Example

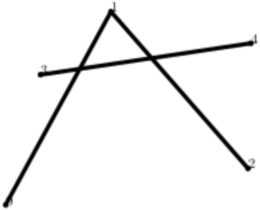


Based on the example from before:

With the edit path $= \{0 \to 0, 1 \to 1, 2 \to 2, 3 \to 3, 4 \to 4, 5 \to \varepsilon, \varepsilon \to 5, \varepsilon \to 6, \varepsilon \to 7\}$



$m_{g_i \times g_j}$ (unpruned)



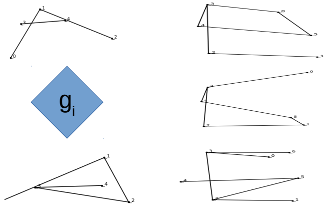
$m_{g_i \times g_j}$ (pruned)

Method

How can we use these matching-graphs to improve the distance calculation between two graphs?

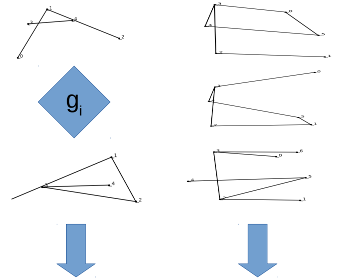
Train

Class A Class E

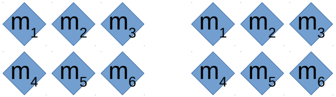


Train

Class A   Class E



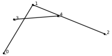
Matching Graphs A   Matching Graphs E



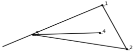
UNIVERSITÄT BERN

Train

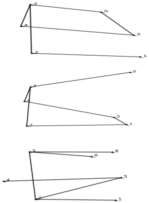
Class A



g_i



Class E



Matching Graphs A

m_1 m_2 m_3

m_4 m_5 m_6

Matching Graphs E

m_1 m_2 m_3

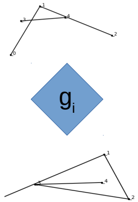
m_4 m_5 m_6

Test

g

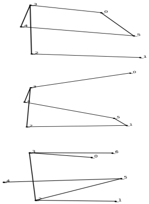
Train

Class A



g_i

Class E



Matching Graphs A

m_1 m_2 m_3

m_4 m_5 m_6

Matching Graphs E

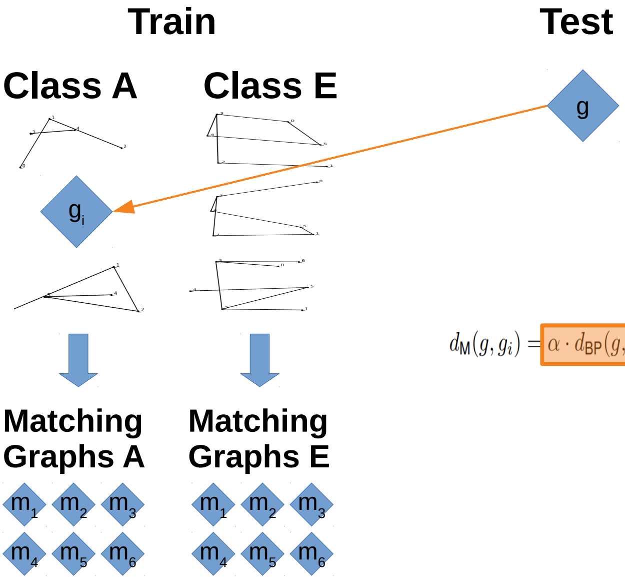
m_1 m_2 m_3

m_4 m_5 m_6

Test

g

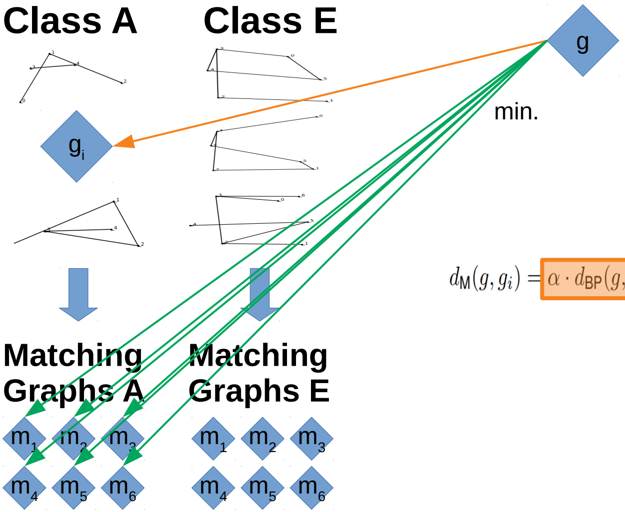
$$d_{\mathsf{M}}(g, g_i) = \alpha \cdot d_{\mathsf{BP}}(g, g_i) + (1 - \alpha) \cdot \min_{m \in \bar{\mathcal{M}}_{\omega_l}} d_{\mathsf{BP}}(g, m)$$



$$d_{\mathsf{M}}(g, g_i) = \alpha \cdot d_{\mathsf{BP}}(g, g_i) + (1-\alpha) \cdot \min_{m \in \bar{\mathcal{M}}_{\omega_l}} d_{\mathsf{BP}}(g, m)$$

UNIVERSITÄT BERN

Train **Test**



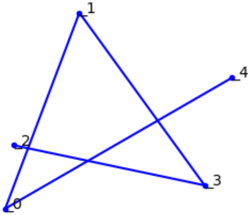
$$d_{\mathsf{M}}(g, g_i) = \alpha \cdot d_{\mathsf{BP}}(g, g_i) + (1 - \alpha) \cdot \min_{m \in \bar{\mathcal{M}}_{\omega_l}} d_{\mathsf{BP}}(g, m)$$

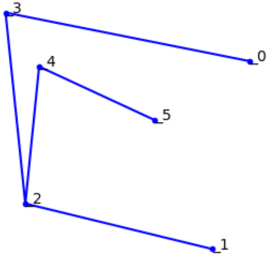
UNIVERSITÄT BERN

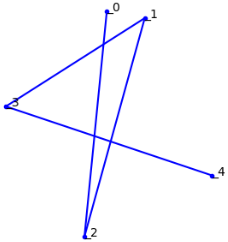
Datasets

Letter

15 Classes: A,E,F,H,I,K,L,M,N,T,V,W,X,Y,Z



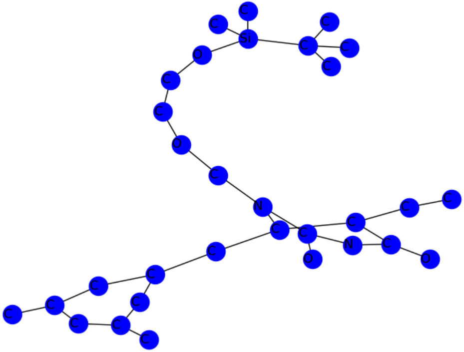


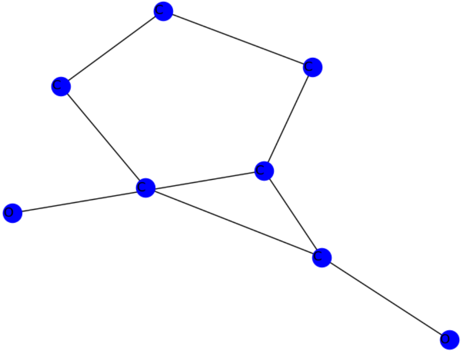


Datasets

Aids

Two classes: confirmed active and confirmed inactive

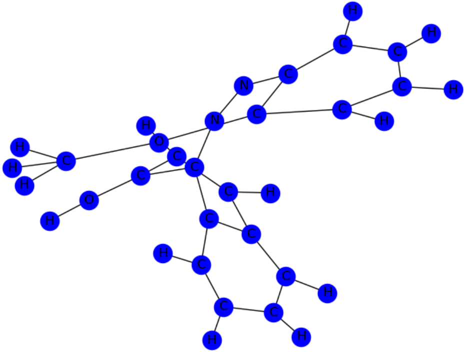


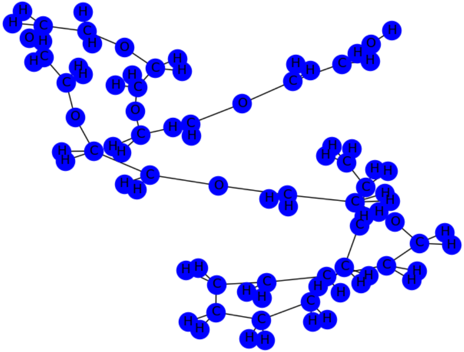


Datasets

Mutagenicity

Two classes: mutagen and non-mutagen





Results

Data Set	k-NN(d_{BP})	k-NN(d_{M}) Unpruned	k-NN(d_{M}) Pruned
Letter	90.5	91.3	93.1 ∘
AIDS	99.0	99.7 ∘	99.7 ∘
Mutagenicity	70.6	70.0	70.5

Conclusion

- Proposal to use matching-graphs, that are pre-computed on training graphs, to improve graph classification.
- These matching-graphs leverage the information provided by the edit path between two graphs
- Initial experiments show promising results.

Future work

- Other graph based matching-graph representations
- Combine the matching-graphs with different classifiers
- Quantitative analysis of the produced matching-graphs