

Master Thesis Seminar Talk

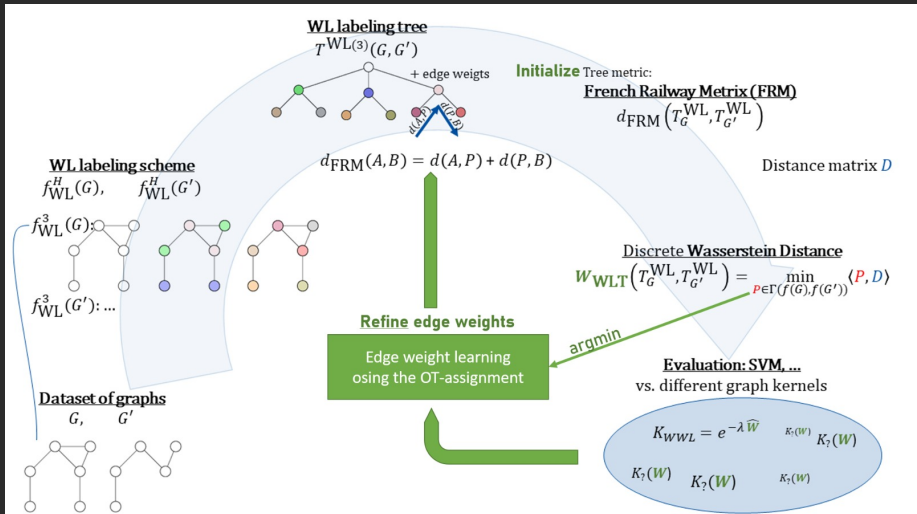
Progress Upade

Fabrice Beaumont

Lab Development and Application of Data Mining and Learning Systems:
Machine Learning and Data Mining

Supervisor: **Dr. Pascal Welke**

19. January 2022



Does the (chosen) iterative weight adjustment
lead to a significant increase in the
expressiveness of the resulting kernel?

1. Given: Graph database $\mathcal{D} = \{(V_i, E_i, \ell_i^V)\}_{i \in [N]}$
Distance d on $\bigcup_{i \in [N]} \text{Range}(\ell_i^V)$ for the FRM
Ground distance d_0 for the Wass. Dist. \mathcal{W}_{WLT}
2. Compute t iterations of Weisfeiler Lehman (WL) labels on \mathcal{D}
3. Construct the WL labeling tree (WLLT) [WL labeling hierarchy]
4. Define edge weights on the WLLT - using a FRM (and d)
5. Define an initial distance between graphs - using \mathcal{W}_{WLT} (and d_0)

1. Given: Graph database $\mathcal{D} = \{(V_i, E_i, \ell_i^V)\}_{i \in [N]}$
Distance d on $\bigcup_{i \in [N]} \text{Range}(\ell_i^V)$ for the FRM
Ground distance d_0 for the Wass. Dist. W_{WLT}
2. Compute t iterations of Weisfeiler Lehman (WL) labels on \mathcal{D}
3. Construct the WL labeling tree (WLLT) [WL labeling hierarchy]
4. Define edge weights on the WLLT - using a FRM (and d)
5. Define an initial distance between graphs - using W_{WLT} (and d_0)

Loop:

- ▶ Use the distances to define a WWL-kernel and classify the database using an SVM
- ▶ Use the mapping chosen by W_{WLT} to identify crucial edge weights.
Refine the edge weights, given the desired closeness of the graphs.

Programming:

- ▶ Construction of the WL-labeling tree (WLLT)
- ▶ Several distance metrics for this WLLT
- ▶ Construction of the WL-label set representation of the graphs
- ▶ Construction of a distance matrix/kernel on the dataset

Still in progress: Several optimizations w.r.t to these implementations.

- ▶ Write an **exposé** to sketch and summarize these research plans
- ▶ Implement the usage of the Wasserstein Distance
- ▶ Chose (several?) update steps or learning methods to adjust the *hot weights*. First:
 - ▶ Constant update with a fixed margin η
- ▶ Implement a **feedback-system**. (An evaluation of the used weights)
- ▶ Literature research

Thank you all for listening.

I will be happy to answer any **questions** and
hear your **comments**.