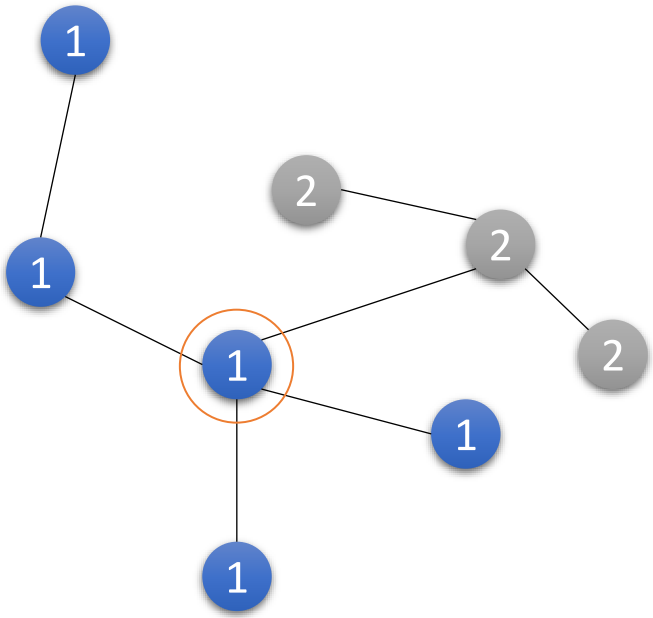
FiGURe: Simple and Efficient Unsupervised Node Representations with Filter Augmentations



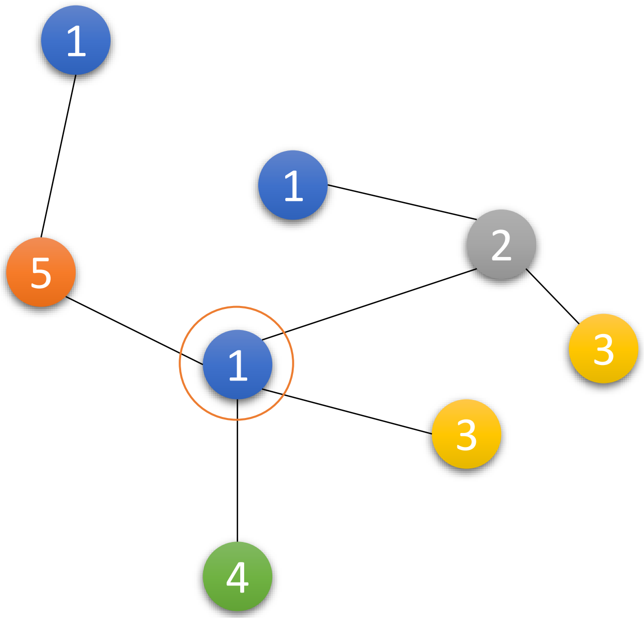


Chanakya Ekbote* Ajinkya Deshpande* Arun Iyer Ram Bairi Sundararajan Sellamanickam

Homophily and Heterophily



Homophily: Majority of the neighbors belong to the same class

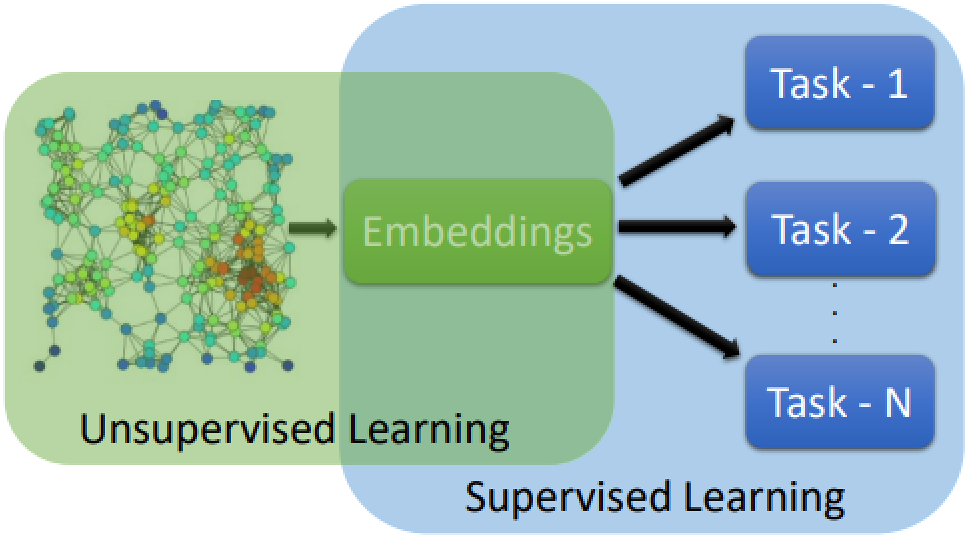


Heterophily: Majority of the neighbors belong to different classes

Problem Setting

Given a graph and node features:

Generate embeddings that: work on tasks with different levels of homophily



GPRGNN (Or where do GCNs fail)?

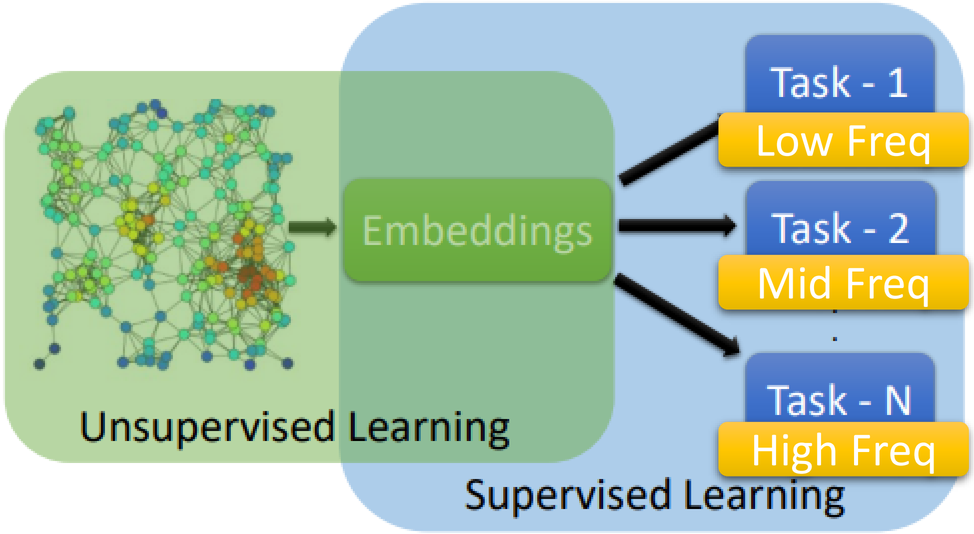
$$Z = \sum_{i=0}^{N} \gamma_i A^i X W$$

$$Z = \sum_{i=0}^{N} \gamma_i F_i X W$$

- Long-range information is not effectively leveraged by conventional GCNs.
- Conventional GCNs face challenges with tasks involving heterophilic graphs as data.
- Fine-tuning of coefficients γ_i is necessary for downstream tasks.

Learn Embeddings Per Filter

Problem

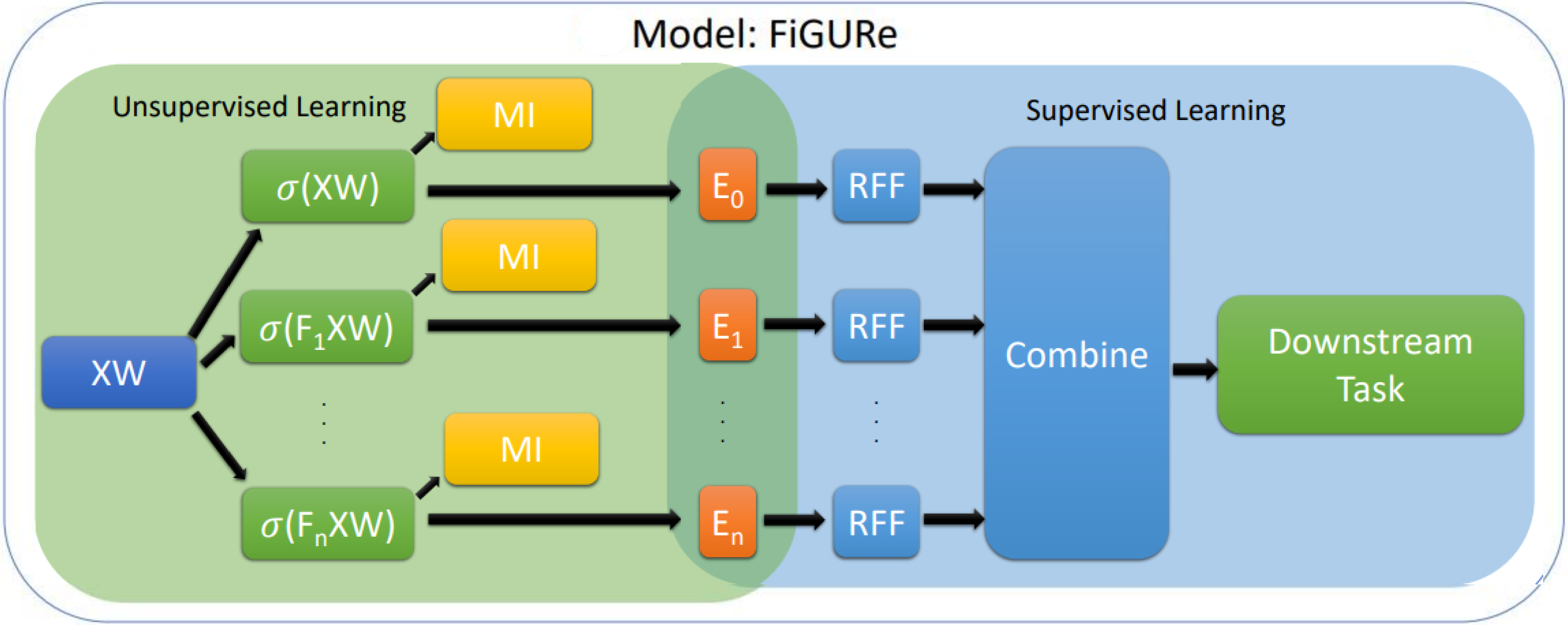


Takeaways

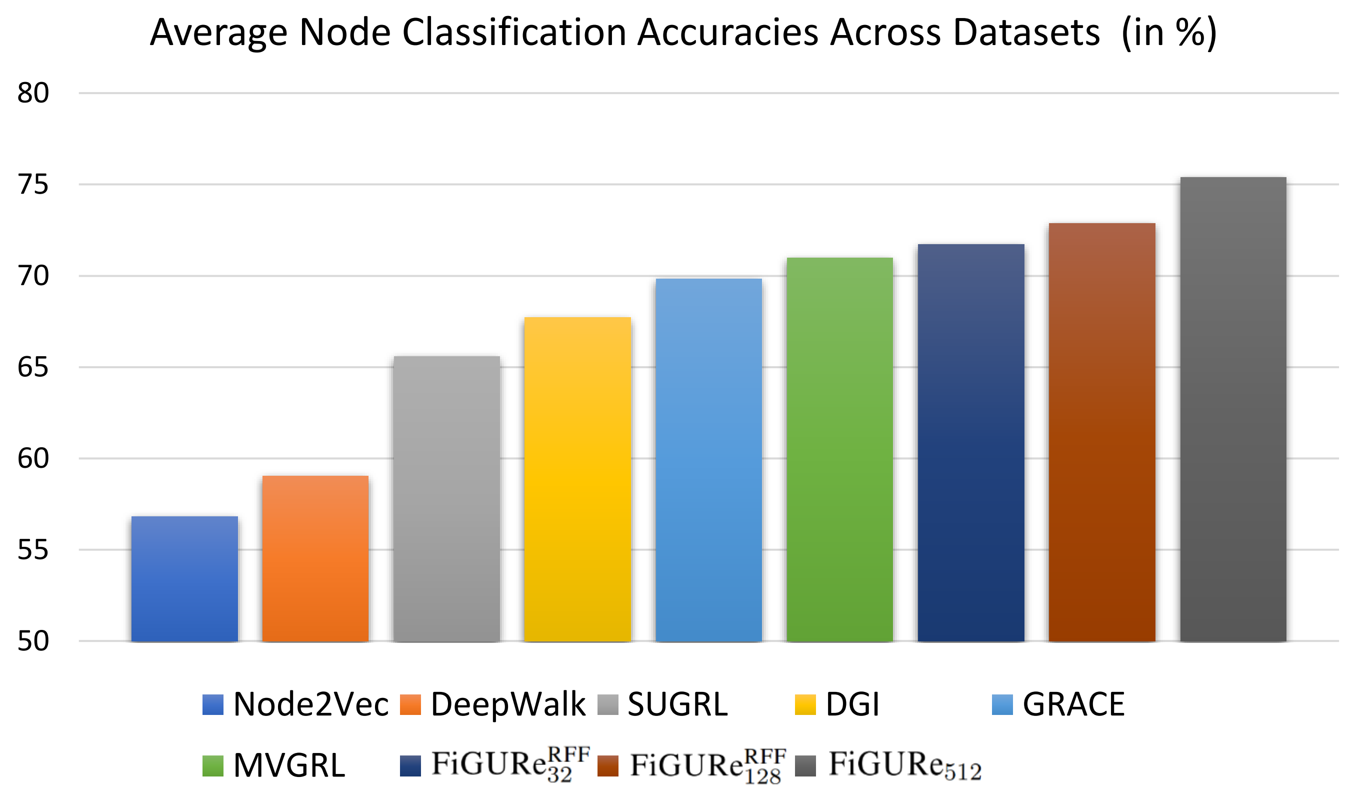
Main Takeaway - 1: Need **to learn embeddings for different filters** that can be **combined in different ways for downstream tasks**.

Main Takeaway - 2: **Storage cost** of multiple large sized embeddings **is huge**.

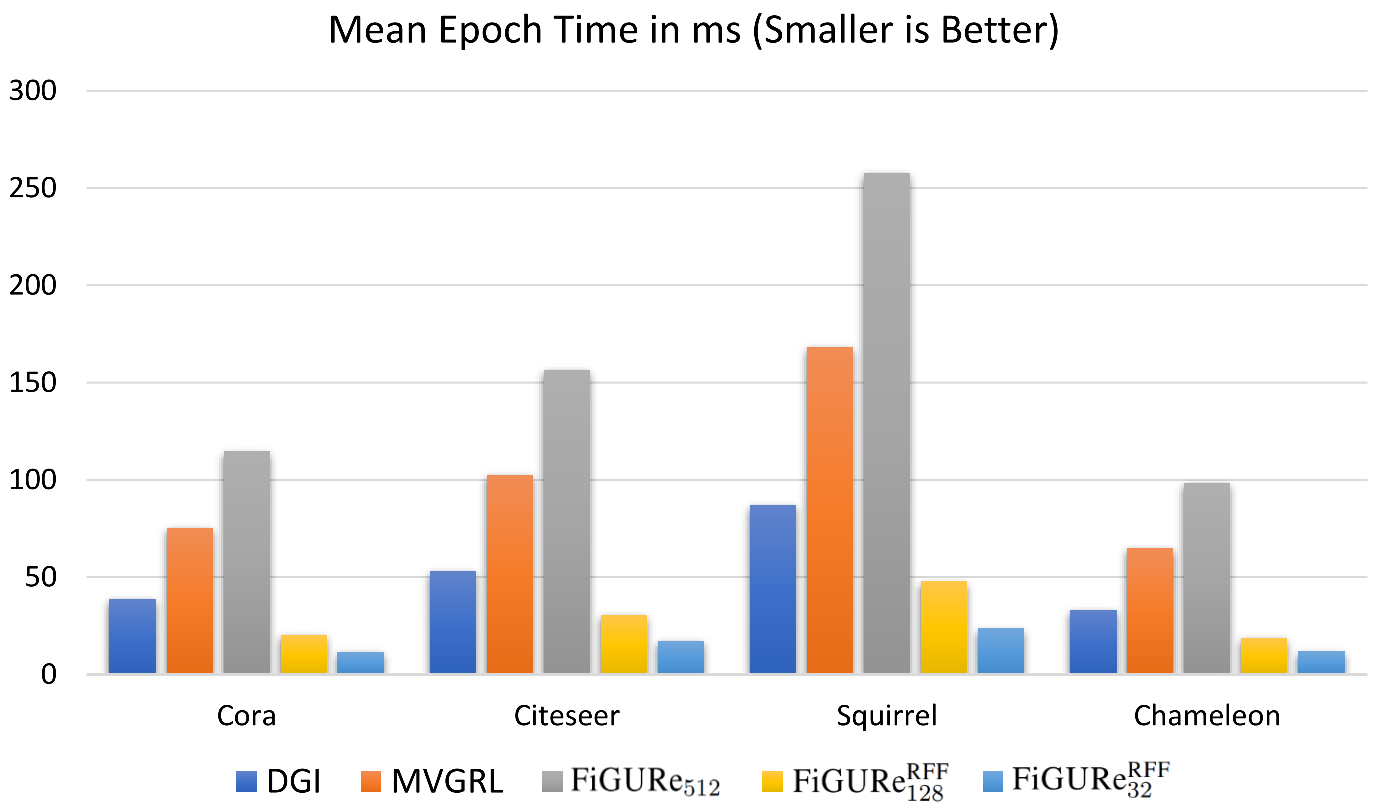
FiGURe


Model: FiGURe
Unsupervised Learning
Supervised Learning
XW
σ(XW)
σ(F1XW)
σ(FnXW)
MI
MI
MI
E0
E1
En
RFF
RFF
RFF
Combine
Downstream Task

Comparison with SoTA Models



RFF – Computational Efficiency



Conclusion

- Enhancing graph contrastive learning with **filter-specific representations**
- Alleviating computational/storage burdens through **low-dimensional representations** and preserving the performance using **RFF**
- Future directions involve
  - Expanding the theoretical analysis of contrastive learning to graphs
  - Investigating linear separability in lower dimensions

Contact

- Correspondence: chanakya.ekbote@epfl.ch (Chanakya Ekbote)

- Link to source code: https://github.com/microsoft/figure