
AutoGL

Automated machine learning for
Graphs

Exploring the Paper: [Automated Machine Learning on Graphs: A Survey](#) by Ziwei Zhang ,
Xin Wang and Wenwu Zhu from Tsinghua University, Beijing, China

Submitted By:
Surabhi Govil(014504984)

What are Graphs:

- Graphs are a structure consisting of objects where some pair of these objects might be related.
- Graphs consists of nodes and edges.
- These can be directed and undirected, multi and hypergraphs, graphs with or without self-edges.
- **Graph Neural Networks** (GNNs) are a common choice to work with graph data.

Graph Machine Learning:

- A branch of machine learning that deals with graph data.
- Application areas include but not limited to recommender systems, particle physics, computational pharmacology/chemistry/biology, traffic prediction, fake news detection, etc.
- Graph Neural Networks (GNNs) are required to implement machine learning to model anything as a graph.

AutoML:

- Process of automating time consuming iterative tasks of machine learning.
- Accelerate the time it takes to get production-ready ML models with great ease and efficiency.
- Helps with feature engineering, hyper parameter tuning, model selection, development and deployment.
- Data Scientists and ML Engineers can concentrate on improving the algorithm rather than worrying about mundane tasks.

Levels of performing AutoML:

- Level 0: No automation. Data scientists code algorithms from scratch.
- Level 1: Use of high-level APIs.
- Level 2: Automatic hyperparameter tuning and model selection.
- Level 3: Automatic feature engineering, feature selection, and data augmentation.
- Level 4: Automatic domain and problem-specific feature engineering, data augmentation, and data integration.
- Level 5: Full automation. No input or guidance required to solve ML problems.

AutoGL:

- AutoML for graph machine learning tasks.
- AutoGL Dataset inherits from PyTorch Geometric to build dataset for graph learning
- AutoGL Solver combines the methods for feature engineering, hyper parameter optimization, model training and ensembling.

Why AutoGL:

- Takes away requirement for manual work and domain expertise required in GNNs for analyzing non-Euclidean data.
- GNNs require tuning many parameters of the graph neural architectures.
- Other AutoML algorithms like Hyper Parameter Optimization (HPO) are not scalable.
- Neural Architecture Search (NAS) can be computationally expensive for larger search spaces.
- AutoGL can deliver high performing models with lesser lines of code.

Conclusion:

- AutoML libraries like AutoSklearn, AutoGluon, NNI, AutoKeras are widely used but cannot be directly applied to graph data.
- AutoGL tends to the uniqueness and complexity of graph tasks.
- AutoGL is open source and easy to use.
- AutoGL is modular thus easily customizable according to needs.
- Eliminates the need for humans inside the experiment loop.

References:

- <https://appen.com/blog/automl/>
- <https://fullstackfeed.com/six-levels-of-automated-machine-learning/>
- <https://arxiv.org/pdf/2103.00742v3.pdf>
- <https://analyticsindiamag.com/meet-autogl-the-first-ever-automl-framework-for-graph-datasets/>
- <https://autogl.readthedocs.io/en/latest/>

Thank You