AutoML in Cybersecurity - Activity 01

1 Purpose

This activity aims to build fundamental knowledge of AutoML principles, theories and develop firsthand experience with open-source AutoML libraries. The activity is divided into the following substacks:

a. Reading and summarizing selected articles on AutoML [1]

b. Explore available open-source AutoML for different machine learning libraries:

2 Reading Papers

Things to do:

- Read and summarize the abstract and conclusion of all the selected papers
- Select 5 papers and provide a detailed summary of those papers.
- Select 1 paper to discuss in details.
- Outline any unclear point or concept.
- Suggest 3-5 articles on AutoML that are not from the selected list of articles.

3 AutoML Libraries

Things to do:

- Required: install and test Hyperopt-Sklearn, Auto-Sklearn, and TPOT for scikit-learn.
- Optional: install and test Auto-Keras and Auto-PYTORCH
- For each library write one or two paragraphs to describe how the library work and the key features of the library
- Suggest other open source AutoML libraries to explore in the future.

[1] the article could be found on `https://github.com/ebinsaad/secautoml`

Useful Resources

- *Automated Machine Learning (AutoML) Libraries for Python*, tutorial by Jason Brownlee, Sept 18, 2020 https://machinelearningmastery.com/automl-libraries-for-python/
- *3 Reasons Why AutoML Won't Replace Data Scientists Yet*, Post by By Marcia Oliveira, a Senior Data Scientist at Skim Technologies, March, 2019 https://www.kdnuggets.com/2019/03/why-automl-wont-replace-data-scientists.html
- *6 Open Source Automated Machine Learning Tools Every Data Scientist Should Know* Post by Shyam Sunder Kumar, August 8, 2020 https://medium.com/analytics-vidhya/6-open-source-automated-machine-learning-tools-every-data-scientist-should-know-49960c1397c9