Polytech network form for PhD Research Grants from the China Scholarship Council

This document describes the PhD subject and supervisor proposed by the French Polytech network of 14 university engineering schools. Please contact the PhD supervisor by email or Skype for further information regarding your application.

Supervisor information	
Family name	TALBI
First name	EL-ghazali
Email	el-ghazali.talbi@univ-lille.fr
Web reference	http://www.lifl.fr/~talbi
Lab name	CRISTAL
Lab web site	https://www.cristal.univ-lille.fr/
Polytech name	Polytech Lille
University name	University of Lille
Country	France

PhD information	
Title	Automatic optimization of hyperparameters in deep learning
Main topics regards to CSC list (3 topics at maximum)	Computer science, Machine learning, deep learning, big data, optimization, evolutionary algorithms.
Required skills in science and	Python programming, optimization, machine

engineering	learning.

Subject description (two pages maximum)

Nowadays there is a lot of excitement around artificial intelligence (AI) with other branches such as machine learning (ML) and Deep Learning. Deep Learning has a huge impact in our lifes. Applications of Deep Learning will rule the world in the near future: autonomous vehicles, langage translation, image recongnition, healthcare, finance, and so on.

Hyperparameter automatic optimization is a big part of Deep Learning. The performance of deep learning models are highly dependent on the choice of the hyperparameters. The automatic search for the optimal values of hyperparameters is of crucial importance in practice and tedious task for many supervised machine learning applications. Hyperparameters are the variables which will determine the network structure (ex. number of Hidden layers and units) and the variables which determine how the network is trained (ex. learning rate, network weight initialization, activation function, momentum, number of epochs, batch size).

The traditional techniques to handle this problem are: manual search, grid search, and random search. The scientific challenges for this project can be described by the following difficulties:

- Size of the search space: deep learning networks are notoriously difficult to set because of the huge number of parameters to configure.

- Expensive objective function: the objective function of the problem is a whole classification, which is highly expensive to evaluate. Hence Bayesian optimization must be applied to handle this issue. It consists in using surrogate models (i.e meta-models, reduced models) such as Gaussian processes (i.e Kriging) to approximate the objective function.

In Bayesian optimization, a surrogate model is first constructed using few evaluation of the objective function. Then the surrogate, which is cheaper to optimize than the original objective, is used to find the next solution to evaluate by applying a criterion to the surrogate (e.g. Expected improvement). Then the surrogate model is improved according to the new generated solutions. It differs from grid search and random search by using past evaluations to select the next solutions to evaluate.

The goal of this project is to:

- Design of a suitable surrogate model for the problem. We will investigate different models such as RBF (Radial Basis Functions), Random Forest, (Deep) Gaussian Processes.

- Design an efficient and robust metaheuristic (e.g. evolutionary) algorithm to finf the best hyperparameters of deep learning for a given application. The designed metaheuristic will be assisted by the surrogate models.

- Validate the proposed methodology on different applications such as image classification and using standard benchmarks of the literature such as MNIST and CIFAR-10. Some other applications can also be investigated such as autonomous and connected electrical vehicles, distributed energy smart grids, smart agriculture and food systems, next generation weather and climate prediction, and smart disaster response.

References

- E-G. Talbi, Metaheuristics: from design to implementation, Wiley 2009.
- I.Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT, 2016.
- Y. Bengio, Y. LeCun, G. Hinton, Deep learning, Nature, 521:436-44, 2010.
- M. Najafabadi, F. Vilanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, E. Muharemagic, Deep learning applications and challenges in big data analytics, Journal of big Data, 2015.
- D. R. Jones, M. Schonlau, W. J. Welch, Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, *13*(4), 455-492, 1998.
- https://machinelearningmastery.com/grid-search-hyperparameters-deep-learning-models-python-keras/