

# Stage de Recherche (6 mois)

## Optimisation et Programmation Génétiques

Contact: Benoît Humez - benoit.humez@sgcib.com - +33 1 42 13 89 93

La programmation génétique (PG) est un domaine prometteur de l'intelligence artificielle. Pour simplifier, la programmation génétique est un programme qui crée des "bons" programmes en s'inspirant de la génétique. A partir de programmes "parents", elle crée des programmes "enfants" par croisements et mutations; ensuite une sélection s'effectue pour garder les meilleurs individus et fabriquer une nouvelle génération.

En restant très prudent sur les attentes que l'on peut avoir, il est intéressant de l'appliquer à la recherche de stratégies d'investissement pour nos clients. C'est dans ce cadre prospectif que nous proposons un stage de recherche de 6 mois au sein du département « Corporate and Investment Banking » de la Société Générale, à Paris La Défense. Les ingrédients pour déterminer les stratégies d'investissement seront les historiques de cours d'actifs financiers, les volumes et les ratios financiers fondamentaux.

En collaboration avec le maître de stage, le candidat devra lors du stage:

- s'approprier le fonctionnement de la PG et arriver à une compréhension fine des principaux concepts (croisements, mutations, introns, "bloat", ...).
- implémenter la PG efficacement et construire une architecture de code appropriée. Il devra notamment se pencher sur les problèmes de performance et sur l'utilisation du parallélisme.
- porter une attention particulière au risque de sur-apprentissage. Il devra développer les moyens de quantifier le sur-apprentissage et réfléchir aux contraintes et limitations à utiliser pour éviter cet écueil.
- donner une interprétation financière aux résultats (est ce un retour à la moyenne ? un suivi de tendance ? quels cas ont été éliminés ? quels sont les filtres trouvés ? ...) afin de valider la stratégie.

- (si le temps le permet) mettre à profit la possibilité d'identifier plusieurs optimums locaux. Avoir des solutions différentes est intéressant en finance: d'une part cela procure plus d'opportunités de placement (la liquidité disponible sur le marché est une contrainte importante), d'autre part cela permet de diversifier le risque. A ce titre, dans la sélection d'une génération, il faudra faire attention à garder les différentes zones intéressantes (« vallées ») et ne pas s'intéresser uniquement à l'optimum absolu. L'idéal étant d'avoir un front de Pareto de programmes optimaux.

Profil recherché pour le candidat:

- autonome, créatif
- appétit pour l'algorithmique
- culture générale en mathématiques appliquées
- bonne connaissance en programmation (C/C++ de préférence) – des connaissances sur le parallélisme et la représentation en arbre sont un plus.
- des connaissances en intelligence artificielle / optimisation sont un plus.
- il n'est pas nécessaire d'avoir des notions en finance

Si l'on agrège le temps passé dans chaque activité, le déroulement ressemblera certainement à:

- 2 mois de lecture/recherche
- 3 mois d'implémentation
- 1 mois de tests

Bibliographie :

- John Koza : Genetic Programming I,II,III,IV
- Deb ; Agrawal ; Pratap ; Meyarivan : A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II