Détection d'anomalies dans l'enregistrement séquentiel des évènements serveurs (logs)

Encadrant : Waïl BENFATMA

L'essor des nouvelles technologies et la digitalisation ont permis une explosion du nombre de données. L'analyse et le traitement de ces données devient alors un enjeu majeur pour rester compétitif et/ou avoir un avantage concurrentiel.
De plus les avancées dans le domaine de l'apprentissage automatique (Machine Learning) permettent d'aller plus loin dans ces analyses et de mettre en place des modèles mathématiques pour, par exemple, détecter un objet dans une image ou déceler une prémisse de pannes sur un système.

Le projet proposé pour cet encadrement concerne le traitement et l'analyse massif de l'enregistrement séquentiel des évènements serveurs (logs). Ceux-ci sont générés de façon quasi-continu et fournissent des informations importantes sur l'état de systèmes.
L'enregistrement affectant un processus particulier (Application, réseau informatique, etc…) permet, dès qu'une erreur survient, de détailler en partie la source du problème et d'en identifier la cause.
Cependant, la volumétrie de ces données est importante et l'analyse de ces logs peut être fastidieuse voire impossible lorsqu'on est submergé par la quantité d'information.

L'objectif de ce projet sera de construire un outil d'analyse de flux de logs continu qui puisse être capable de détecter tout événement susceptible de provoquer une panne serveur.