

COMMUNIQUE DE PRESSE – Juillet 2014

Le Big Data arrive dans l'édition : Short Edition développe l'algorithme de prédiction de la qualité littéraire

Short Édition, l'éditeur communautaire de la littérature courte, a lancé depuis le début de l'année 2014 un programme de recherche fondé sur la **technologie du Machine Learning**, visant, en particulier, à développer **un système intelligent qui assistera le Comité Editorial pour l'évaluation de la qualité littéraire d'un texte.**

Des évaluations sur 25 000 œuvres : la base d'apprentissage de l'algorithme

Pour s'assurer de la qualité des œuvres publiées et mises en avant sur le site, Short Édition a mis en place un comité éditorial de grands lecteurs qui évaluent les œuvres selon un protocole encadré quasi-scientifique.

Aujourd'hui, ce sont plus de 25 000 œuvres qui ont, chacune, été évaluées par 5 à 10 membres du Comité Éditorial pour les besoins de publication. « *Ces évaluations représentent une formidable base d'apprentissage pour un algorithme d'intelligence artificielle* » précise avec enthousiasme **Quentin Pleplé, Data-Scientist, CTO de la start-up grenobloise.**

Le principe d'apprentissage automatique sur du Big Data repose sur un modèle éprouvé et reconnu : des individus réalisent une tâche définie sur des données, leur travail sert alors de modèle à un algorithme qui apprend à reproduire la tâche. L'algorithme peut ensuite exécuter la tâche sans intervention humaine, avec un taux d'erreur que l'on mesure. L'algorithme est finalement utilisé en conditions réelles lorsque le taux d'erreur est jugé acceptable.

L'algorithme prédictif issu du programme de R&D de Short Édition sera capable de mimer le travail d'un membre du comité éditorial et de donner **une probabilité sur la qualité littéraire de nouveaux textes.**

« *La force et la légitimité de Short Edition, c'est que nous sommes les seuls à posséder ce type de jeu de données : un grand nombre d'œuvres variées dont la qualité littéraire est évaluée rigoureusement par un nombre important de lecteurs experts profilés* » précise le jeune polytechnicien pilote du programme de recherche, associé fondateur de Short Edition.

Un grand intérêt pour Short Edition mais aussi pour l'édition et pour la presse

Short Édition souligne que la machine ne remplacera pas le cerveau humain, mais qu'elle l'assistera pour le rendre plus efficace : **l'algorithme sera utilisé comme une aide à la décision** mais en aucun cas ne pourra se substituer à l'avis du comité éditorial.

La puissance du Machine Learning et du Data Mining permettra à Short Edition de développer son modèle communautaire d'une part en décuplant le travail du Comité éditorial par le pré-classement et la pré-évaluation des œuvres soumises, et d'autre part en

mettant plus facilement en avant les meilleures œuvres sur le site, sur les applis et dans les boîtes mails des lecteurs.

Cette innovation aura aussi un grand intérêt pour les acteurs de l'édition (pré-évaluation des manuscrits) et pour les acteurs de la presse (médiats ou contenus participatifs, nouveau modèle économique avec le digital) qui vont être confrontés à la problématique de gestion d'un très important flux entrant de textes et de leur traitement.

Un programme de recherche soutenu par de multiples acteurs

Le programme de recherche de Short Édition est réalisé en collaboration avec **trois laboratoires spécialisés en intelligence artificielle** : le LIMSI adossé à l'Université Paris Sud (Orsay), le LIRIS adossé à l'Université Lyon I et le LIG adossé à l'INRIA et au CNRS.

Ce programme ambitieux est soutenu par **BPI France** (ex OSEO) à partir de 2014. La start-up est par ailleurs accompagnée par **ProbaYes**, société spécialisée en apprentissage automatique et en analyse prédictive, basée à Grenoble.



ProbaYes
Mastering Uncertainty

Le DEFT, compétition inter-laboratoires de Big Data, a choisi Short Édition

Le DEFT (Défi En Fouille de Textes, deft.limsi.fr) est une compétition annuelle organisée par le laboratoire LIMSI (adossé à l'Université Paris-Sud) de fouille de textes Big Data, entre une dizaine de labos d'universités francophones du monde entier et d'entreprises multinationales (Orange, IBM...) qui sont en compétition sur une tâche de fouille de texte.

En 2014, pour la dixième édition de leur compétition, les organisateurs du DEFT ont choisi comme thème la problématique de recherche de Short Edition. Les concurrents ont eu 14 semaines pour créer, à partir d'une partie des données de Short Edition, un prototype de l'algorithme prédisant la qualité littéraire.

A propos de Short Édition

Short Édition est un éditeur de littérature courte (**tout ce qui se lit d'un trait, en moins de 20 mn** : Nouvelles, BD courtes, Poèmes et Micro-nouvelles) et anime une communauté de lecteurs et d'auteurs qui osent aimer la littérature courte.

Short Édition compte aujourd'hui **82 000 lecteurs abonnés** et a généré **plus de 3,1 millions de lectures d'œuvres** et a donné sa chance à **plus de 4 000 auteurs** sur sa plateforme participative short-edition.com et sur ses applis iOS et Android pour smartphones et tablettes.

Short Édition, qui a reçu le Trophée de l'innovation catégorie Culture au Printemps numérique 2013 et a été lauréat de la Fête nationale des Services 2014, vient de faire entrer **dans son capital** (avril 2014) **le Groupe Express Roularta** qui édite notamment les supports *L'Express*, *L'Etudiant*, *Lire* et *L'Expansion*.