

## L'open data et l'open source, des soutiens nécessaires à une justice prédictive fiable ?

Michaël Benesty  
*Avocat fiscaliste, Data Scientist*

**Abstract.** Une expérimentation menée sur des décisions de cours d'appel administratives françaises avec un algorithme de machine learning a permis d'établir que la confiance que l'on placerait dans des outils de justice prédictive implique, au cours du processus, que les calculs et les caractéristiques du modèle de prédiction soient visibles et compréhensibles par le juriste, avocat ou magistrat, qui les emploient. Cela se traduit par une obligation de transparence sur les algorithmes, qui est garanti par la libération en open source de la solution autant que par une vigilance et une expertise juridique sur les données mobilisées.

**Keywords:** Jurisprudence, Open Source, Open Process, Machine Learning, Analyse Prédictive

Depuis le début des années 2000, un portail<sup>1</sup> mis en place par l'État français permet aux justiciables d'accéder gratuitement à un fonds non exhaustif de décisions des principaux tribunaux et cours du pays. En 2014, ces mêmes documents ont été mis à disposition de tous sous forme de fichiers directement exploitables, permettant à ceux qui se sentent inspirés de les exploiter pour leur appliquer, par exemple, des traitements informatiques. Enfin, la loi pour une République Numérique<sup>2</sup> votée fin 2016 a rendu officielle l'ouverture de presque toutes les décisions de justice.

En parallèle du mouvement de l'open data juridique, les algorithmes de machine learning (souvent désignés par l'expression "Intelligence Artificielle" dans la presse générique) connaissent un très fort regain d'intérêt depuis le début des années 2010. Ces algorithmes sont basés sur des approches probabilistes et se veulent plus souples que les autres en étant capables de s'adapter aux données auxquels ils s'appliquent. Ils sont souvent utilisés pour extraire une ou plusieurs caractéristiques pour chaque observation d'un jeu de données. Par exemple, nombre d'algorithmes de

---

<sup>1</sup> Legifrance, service public de diffusion du droit par l'Internet, ouvert en 1998, régi par le décret 2002-1064 du 7 août 2002

<https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000413818>

<sup>2</sup> LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique. En ligne : <https://www.legifrance.gouv.fr/affichLoiPubliee.do?idDocument=JORFDOLE000031589829&type=general&legislature=14>

machine learning ont pour objet de classer une observation dans une liste de catégories prédéfinies.

Appliqué au texte d'une décision de justice, un algorithme de machine learning peut par exemple servir à classer une décision en fonction de son thème, ou dégager parmi un ensemble de décisions celles pour lesquelles le demandeur a eu gain de cause.

Ces quelques tâches simples peuvent être reformulées en tâches plus audacieuses. Par exemple, apprendre à reconnaître les décisions pour lesquelles le demandeur obtient gain de cause, sans avoir accès au texte du dispositif de la décision, revient à pousser l'algorithme à apprendre si la combinaison des arguments développés dans la décision permet de gagner le contentieux.

Lors d'un important colloque portant sur l'open data juridique et organisé par la Cour suprême française de l'ordre judiciaire (14 octobre 2016, Cour de cassation, La jurisprudence dans le mouvement de l'open data<sup>3</sup>), le thème de la "justice prédictive" a été évoqué par un grand nombre de professionnels du droit, en particulier par les magistrats.

La justice prédictive consiste à utiliser des algorithmes de machine learning pour "prédire" l'issue d'un procès ou détecter les arguments qui sont les plus déterminants.

Malgré le nombre d'écrits traitant du sujet, des points essentiels restent peu souvent abordés. En particulier, la transparence dans la mise en oeuvre des algorithmes et la compréhension des résultats est rarement discutée. Pourtant la question est cruciale. Il est légitime de penser que les résultats générés par des solutions opaques (en particulier ceux provenant de sociétés commerciales) ne satisfont pas l'exigence de transparence et limite de ce fait la confiance que l'utilisateur peut leur accorder.

Donner accès au code source et aux données est une manière de permettre aux utilisateurs (magistrats, avocats, ou justiciables) de mesurer objectivement et de façon indépendante la qualité et la pertinence du résultat généré. Par ailleurs, certains aspects clés de l'algorithme liés à des questions juridiques peuvent bénéficier de contributions extérieures comme c'est le cas dans l'industrie du logiciel en général avec le mouvement open source.

Pour faire un parallèle avec le droit, l'algorithme prédictif opaque est comme un juge qui motiverait ses décisions en citant des lois auxquelles personne d'autre que lui n'aurait accès.

---

<sup>3</sup> Lien officiel vers le site du colloque

[https://www.courdecassation.fr/venements\\_23/colloques\\_4/2016\\_7423/open\\_data\\_35038.html](https://www.courdecassation.fr/venements_23/colloques_4/2016_7423/open_data_35038.html)

C'est un vrai problème et il ne nous semble pas raisonnable qu'un professionnel du droit puisse un jour utiliser de quelque manière que ce soit une "prédiction" sur laquelle il n'aurait aucun moyen de contrôle. Ce serait similaire à un avocat qui cite dans ses conclusions des arrêts qu'il n'a jamais lus. Cette question touche directement la déontologie et la qualité des travaux rendus aux justiciables.

Il est possible de résumer les relations entre la justice prédictive et la transparence dans les données et les techniques de la sorte :

- l'ouverture des données juridiques par l'État (*open data*) permet la faisabilité technique d'un projet de justice prédictive ;
- l'ouverture du code source par l'auteur du projet (*open source*) donne à l'utilisateur des prédictions les moyens d'avoir confiance dans les résultats générés.

Le présent écrit est accompagné d'un algorithme de justice prédictive en open source au travers du projet OPJ (pour *Open Predictive Justice*).

L'algorithme est appliqué à la délicate question des mesures d'éloignement des étrangers. Il s'agit du contentieux le plus volumineux en nombre de décisions en droit administratif français et certainement l'un des plus graves dans ses conséquences.

L'algorithme présenté permet de prédire l'issue de ce contentieux à hauteur de 87,2 %. Il nous a entre autres permis de découvrir que faire figurer que l'étranger bénéficie de l'aide juridictionnelle (avocat payé par l'État) augmente sensiblement la probabilité de voir la décision d'éloignement annulée. En revanche, l'analyse montre que connaître l'identité du juge n'aide pas de façon significative à prédire l'issue du contentieux.

Le projet OPJ inclut le code source ainsi que le jeu de données associé (plus de 18 000 décisions de justice récentes annotées). Le projet est accessible sur un dépôt Github<sup>4</sup>. Nous plaçons le code sous licence permissive Apache 2.0 de façon à ce que toute personne puisse le réutiliser de la façon dont elle le souhaite.

L'approche choisie dans ce projet est applicable à toutes les questions de droit, qu'il s'agisse, par exemple, de fiscalité ou d'indemnités de licenciement. Elle est aussi transposable aux décisions rendues dans les droits d'autres pays.

Dans cet écrit nous expliquons la démarche que nous avons utilisée de manière à ce qu'elle puisse être répliquée par tout juriste sur toutes les questions de droit.

---

<sup>4</sup> Le projet est accessible ici : <https://github.com/SupraLegem/OPJ>

## 1. Une Définition du Machine Learning

Il s'agit d'une expression vague qui regroupe toute une famille d'algorithmes qui ont en commun d'apprendre par eux-mêmes en observant des données. Ces algorithmes sont inspirés de différentes sciences et notamment des statistiques.

En matière de justice prédictive, c'est la sous-famille des algorithmes supervisés qui est utilisée. Ces algorithmes sont alimentés par des données (ex : des décisions de justice) et apprennent à prédire un résultat (ex : l'issue d'un contentieux).

L'essence du fonctionnement de ces algorithmes est de chercher dans les différentes propriétés des données celles qui sont le mieux corrélées avec le résultat. Par exemple, on peut imaginer que l'algorithme cherche dans chaque décision de justice les arguments soulevés qui sont les plus à même d'obtenir une victoire dans un contentieux.

Grâce à ce procédé, l'algorithme apprend à imiter les données, donc ici, imiter la façon qu'ont les juges de trancher les contentieux au regard des arguments avancés par les parties.

## 2. Comment Mesurer la Qualité des Prédications d'un Algorithme de Justice Prédictive ?

La façon la plus évidente de procéder à cette mesure est de fournir des données à l'algorithme et de mesurer la proportion des décisions dont il prédit correctement l'issue.

Dans le cas d'espèce dont il est question ici, les données sont des descriptions du contenu des décisions. On indique par exemple si le demandeur est une personne physique, si tel numéro d'article est cité ou si tel nom de pays apparaît. Ainsi, l'algorithme comprendra peut être que les décisions qui mentionnent Algérie et l'accord franco-algérien modifié du 27 décembre 1969 ont plus ou moins de chance d'être rejetés comparée à une argumentation basée sur le regroupement familial (article 8 de la CEDH).

L'expérience est menée sur plus de 18 000 arrêts de cours administratives d'appel (2012 à 2015) relatifs au contentieux des mesures d'éloignement des étrangers (Obligation de Quitter le Territoire Français : OQTF ci-après). Certaines décisions sont exclues, en particulier celles se prononçant explicitement sur l'aide juridictionnelle.

Ci-après quand nous parlerons de demandeur personne physique il faudra donc comprendre que c'est un contentieux qui oppose un étranger à l'administration et que l'étranger demande à rester en France. Si c'est l'administration qui est en demande, c'est en général que l'administration a tenté d'éloigner l'étranger mais que le juge du premier degré a annulé la mesure d'éloignement.

L'algorithme apprend à "prédire" s'il y a eu un rejet de la demande ou non. Comme il est de coutume en machine learning, nous mesurerons les résultats sur un jeu de données que l'algorithme n'aura pas vues pendant la phase d'apprentissage.

### 3. Choisir un Algorithme

Il existe différentes familles d'algorithmes supervisés. Pour ce projet, nous utilisons [XGBoost](#)<sup>5</sup>. Cet algorithme :

- représente l'état de l'art en apprentissage de données structurées ;
- est utilisé en production, entre autres, chez plusieurs des [GAFA](#) ;
- produit un modèle interprétable ;
- est disponible en open source sous licence permissive Apache 2.0.

Il existe d'autres choix qui sont susceptibles de produire des résultats similaires voire meilleurs. Cependant XGBoost est réputé pour produire des résultats au minimum proches de l'état de l'art et cela sans chercher à optimiser sa configuration<sup>6</sup>.

### 4. Comment avons-nous extrait les descriptions des arrêts ?

Pour fonctionner, l'algorithme de machine learning ne peut pas être appliqué directement au texte d'une décision. Le texte écrit par un humain n'est pas suffisamment structuré et contient souvent trop d'ambiguïtés pour être facilement "compris" par une machine. Le programme informatique a besoin que le texte soit transformé dans un format qui lui est compréhensible. De nombreuses approches sont envisageables, la plus simple est de présenter chaque texte sous forme d'un grand tableau où les colonnes seraient les informations contenues. Dans le cadre des décisions de justice, l'information dont il est question pourrait être la présence ou non d'un argument ou la citation ou non d'une loi, etc.

Le fonctionnement de l'algorithme repose sur l'extraction de certains éléments contenus dans les décisions de justice :

- la description des faits ;
- la description des arguments utilisés par les parties ;
- les numéros d'articles présents dans la décision ;
- la nature des parties (personne physique ou administration) ;
- le nom du magistrat qui a présidé la chambre qui a rendu la décision.

---

<sup>5</sup> <https://github.com/dmlc/xgboost>

<sup>6</sup> L'auteur de cet article est l'un des développeurs d'[XGBoost](#).

Cette étape est spécifique à chaque question de droit traitée. Pour extraire ces éléments nous avons simplement cherché des chaînes de caractères dans les décisions.

Le texte a été extrait des décisions en open data, mis en minuscule et les retours à la ligne sont indiqués par la mention @L. Les points et les virgules ont été retirés. Ces traitements sont standards en machine learning appliqué au langage.

Toutes les mentions qui sont présentes moins de 200 fois sont retirées (il s'agit d'éviter le bruit dans les données). Plus de 29 000 colonnes ont été générées en combinant le contenu des colonnes existantes.

N'étant pas experts en droit des étrangers, nous sommes allés sur le site du [GISTI](#) (Groupe d'information et de soutien des immigré·e·s) et nous avons consulté le guide "[Contester une obligation de quitter le territoire français, la procédure !](#)"<sup>7</sup> qui présente sur dix pages des arguments juridiques à opposer à l'administration pour faire annuler une mesure d'éloignement. Pour chaque argument présenté, nous avons recherché la façon dont il est rédigé dans les décisions et, via les statistiques calculées par le site, nous avons vérifié que la chaîne de caractères avait un effet sur les pourcentages de rejets de chaque cour administrative d'appel. En fonction de l'élément, nous cherchons sa présence dans tout le texte de la décision ou seulement dans les demandes.

Concernant les numéros d'article, l'extraction est faite automatiquement. La recherche des numéros pouvant donner une information sur l'issue du contentieux (ex : article L. 761-1 du Code de Justice Administrative concernant la condamnation aux dépens) est limitée aux demandes. L'approche est simple et nous permet de garantir que les éléments extraits ont du sens. Une approche reposant exclusivement sur des extractions automatisées depuis le texte est possible mais n'offrirait pas le même degré de contrôle, de transparence et d'interprétabilité.

Il y a un compromis à trouver entre coût de mise en place de l'algorithme (en temps notamment) et qualité des résultats.

Ce choix emporte un certain nombre de conséquences. Il est possible d'imaginer que des questions de droit moins importantes (en volume et en enjeux par exemple) soient traitées davantage de façon automatisée. Cette question n'est pas spécifique à la justice prédictive mais concerne le machine learning en général. L'approche open source permet de poser la question de façon transparente (les procédés utilisés sont connus de tous) et permet à plusieurs personnes d'y répondre de différentes manières.

---

<sup>7</sup> <http://www.gisti.org/spip.php?article2743>

## 5. Quels sont les Résultats ?

Cet algorithme est capable de prévoir correctement 87 % des rejets en se basant sur des extractions depuis le texte de plusieurs milliers de décisions. Les paramètres de l'algorithme n'ont pas été optimisés et nous pensons qu'il n'est pas difficile d'obtenir une amélioration significative des résultats. Pour information, le taux de rejet moyen dans ce jeu de données est de 75 %.

[XGBoost](#) a des fonctions avancées qui permettent d'étudier le modèle qu'il a généré sur la base des données qui lui ont été fournies. Il est ainsi possible de connaître les éléments qui sont les plus déterminants pour la prédiction, c'est-à-dire les arguments ou les faits les plus corrélés à l'issue du contentieux.

Plusieurs types d'analyses sont possibles :

- identifier les caractéristiques les plus décisives pour déterminer l'issue d'un contentieux ;
- identifier les combinaisons (faits + arguments) les plus pertinentes ;
- calculer les pourcentages de réussite pour une combinaison donnée.

Dans le présent article, nous ne nous intéressons qu'au premier point.

Nous avons listé les caractéristiques qui ont été utilisées par l'algorithme pour prédire le dispositif. Les deux plus importantes sont la nature du demandeur et du défendeur. Les étrangers ont en moyenne une forte probabilité (de l'ordre de 80%) de se faire rejeter leur demande d'annulation d'une mesure d'éloignement.

Un grand nombre de caractéristiques qui remontent ont trait au type de moyens soulevés. En particulier, savoir si une défense fondée sur le regroupement familial est décisive pour la prédiction. C'est un moyen bien connu et ce n'est donc pas une surprise.

La caractéristique suivante concerne l'insuffisance de motivation de l'OQTF lorsque le demandeur est l'administration. En faisant quelques recherches sur les bases de jurisprudences, on constate que lorsque l'administration est en demande et que la décision traite d'insuffisance de motivation de l'OQTF, l'administration gagne 90 % du temps (25 points de plus que la moyenne) ! Cette solution est conforme aux dispositions de l'article L. 511-1 du CESDA et, de la lecture des décisions, il semble que quelques magistrats de tribunaux administratifs relevant principalement des cours administratives d'appel de Paris et de Lyon n'en tiennent pas compte. En appel ces jugements sont infirmés. L'enseignement que les avocats doivent en tirer, c'est que ce moyen ne fonctionne pas en appel. Notons tout de même que pratiquement 1 000 décisions du jeu de données y font référence...

La mention de l'aide juridictionnelle (avocat payé par l'État) lorsque le demandeur est une personne privée est très importante pour l'algorithme.

Rappelons que les décisions se prononçant sur l'aide juridictionnelle ont été retirées du jeu de données. On pourrait penser que les décisions mentionnant l'Aide juridictionnelle (AJ) présentent un taux de rejet moyen plus haut car les avocats qui la recevront sont objectivement mal rémunérés par l'État.

Bien que la preuve de la résidence en France ne soit pas exigée pour obtenir l'AJ en cas d'OQTF, il y a des obligations documentaires à satisfaire. Il se pourrait que cela donne au juge une certaine confiance dans le dossier et aboutisse à un taux de rejet un plus bas. Une possible conséquence serait d'inviter les conseils à bien insister sur ce point dans leurs conclusions.

Il y a 1620 caractéristiques remontées.

## **6. Connaître l'Identité de Son Juge Aide t-il à Prédire la Décision en Oqtf ?**

La technique utilisée en machine learning pour déterminer si une caractéristique est critique ou non pour un algorithme s'appelle l'ablation. Il s'agit simplement de comparer le score de prédiction avant et après le retrait de l'information dont on souhaite mesurer l'importance.

Nous avons donc retiré le nom des magistrats des descriptions des décisions. Dans ce contexte, la variation du taux de prédiction est d'un demi-point.

Si certains magistrats ont des taux de rejet en OQTF très élevés par rapport à la moyenne (voire proches de 100%)<sup>8</sup>, ils sont peu nombreux. C'est pour cela que la connaissance du nom du magistrat n'aide pas de façon statistiquement significative l'algorithme lorsque l'on considère le jeu de données dans son ensemble. Cependant, une étude du modèle généré montre que l'algorithme a trouvé de façon autonome les juges qui ont des taux de rejet particulièrement élevés et il s'en sert dans ses prédictions.

## **7. Une Nécessaire Transparence en Justice Prédicative**

La transparence est l'objet même de cet article. Sans transparence il n'est pas possible de mener une analyse critique du résultat. Sans analyse critique, il est impossible de savoir si un algorithme est fiable ou si l'approche est biaisée (même par une erreur faite de bonne foi).

Pourtant, la transparence pourrait être difficile à obtenir de la part de sociétés commerciales.

---

<sup>8</sup> <https://medium.com/@supralegem/l-impartialit%C3%A9-de-certains-juges-mise-%C3%A0-mal-par-l-intelligence-artificielle-ee089170ddd3#.thw885ag5>



Il faut savoir que dans l'industrie du machine learning, tous les grands acteurs (Google, Facebook, Stanford University, etc.) mettent en open source leurs principaux algorithmes (avec quelques exceptions) et publient les résultats de leurs recherches dans des dépôts gratuits comme [arxiv.org](https://arxiv.org) (la référence mondiale, maintenue par l'Université de Cornell) ou [HAL](https://hal.archives-ouvertes.fr/) (la référence en France, maintenue par le CNRS). Il est communément admis que cette pratique a permis les évolutions fascinantes en machine learning auxquelles nous assistons.

Partager le code source n'est pas un problème pour ces acteurs, car ils possèdent des quantités de données brutes que les autres acteurs n'ont pas. Ils conservent donc un avantage concurrentiel malgré l'ouverture de leur code.

En France, cet avantage via l'accès aux données brutes s'annule car toutes les décisions seront progressivement libérées en open data. Cela inclut même des bases antérieures à la Loi pour une République Numérique du 7 octobre 2016 comme Jurica, c'est-à-dire 100 % de la jurisprudence des cours d'appel judiciaires depuis 2007. On disposera d'1,5 million de décisions par an issues de l'ordre judiciaire selon la Cour de cassation. Seules les décisions des tribunaux de commerce ne sont pas dans le champ du dispositif en matière civile.

Nous estimons donc qu'il sera difficile aux sociétés commerciales en France d'ouvrir leurs algorithmes car elles n'auraient alors plus aucun avantage concurrentiel qui justifierait que leurs clients payent un abonnement. Pourtant, l'ouverture du code source est un moyen de permettre la transparence nécessaire pour comprendre d'où vient le résultat.

## **8. Est-il Pertinent d'Apprendre à Prédire l'Issue d'un Contentieux en Analysant des Décisions de Justice ?**

Les conclusions d'un avocat soutenant les prétentions de son client font souvent une dizaine de pages (annexes exclues). Dans de nombreux contentieux, il y a deux parties dans un procès. Chaque conclusion propose une certaine vision de la réalité qui de manière attendue ne concorde pas avec la conclusion de la partie adverse. La version présentée est souvent très "favorable" aux intérêts de l'une ou l'autre partie.

Seuls les éléments jugés pertinents à la compréhension des motifs et du dispositif sont rappelés dans la décision publiée. Dans les contentieux de masse (comme l'OQTF, le licenciement, etc.) cette tendance est encore plus forte (les décisions de cette matière sont très semblables les unes aux autres).

Les décisions de justice rendues en France sont connues pour être faiblement motivées. Elles tiennent souvent sur une ou deux pages (hors

en-tête, etc.). En comparaison, aux États-Unis, [une décision rendue par une cour fédérale fait en moyenne 9,1 pages](#)<sup>9</sup>.

Le problème, c'est que les éléments contenus dans les décisions sont par essence corrélés au dispositif. Une partie des moyens soulevés par un avocat ne sont pas vraiment examinés dans la décision (ils l'ont bien entendu été par le magistrat mais celui-ci ne va pas tout reprendre dans le texte de la décision). Parfois le moyen est surabondant (un autre moyen était plus efficace, ce n'est pas la peine d'étudier les autres), d'autres fois un problème sur la forme empêche l'examen sur le fond, etc. C'est tellement courant que l'absence de motivation est un des principaux motifs de cassation.

Bien entendu, il ne s'agit pas ici de critiquer la manière dont les décisions sont écrites, mais uniquement de comprendre ce qui manque à ce document pour permettre à un algorithme prédictif de fonctionner efficacement.

Or, l'objet d'un algorithme de machine learning est de rechercher la corrélation entre les données et le résultat à prédire. Le simple fait que les données contenues dans les décisions sont par essence corrélées avec le dispositif introduit un [biais de sélection](#) qu'il nous semble difficile de mesurer. Pour éliminer ce biais, il nous faudrait accéder à des données extérieures à la décision comme les pièces du dossier (ces données ne sont pas en open data). La quantité des décisions utilisées pour chaque question de droit traitée permet certainement, dans une certaine mesure, de limiter ce biais (sans mesure objective pour comparer, on ne peut que spéculer).

Les informations latentes dans la décision présentent une autre difficulté. Par exemple, en matière de licenciement d'un manager, est-il pertinent de comparer le sort réservé par un conseil des prud'hommes à un manager d'un restaurant McDonald's avec celui d'un manager d'un cabinet de consultants comme Deloitte ? Comment l'algorithme peut-il comprendre que ces situations ne sont pas comparables si une partie des décisions (notamment celles qui tiennent sur une demi-page A4) ne mentionne pas explicitement la différence de situation (que nous, humains, devinons au travers du nom de l'employeur) ?

Un autre point à considérer est qu'il est possible qu'en fonction du dispositif le juge choisisse une rédaction spécifique. Il est par exemple envisageable que les juges utilisent l'expression "atteinte disproportionnée" lorsqu'ils veulent expliquer que dans le cas jugé il n'y en a pas. Si tel est le cas, il faudra en tenir compte pendant la phase d'extraction des caractéristiques des données.

---

<sup>9</sup> <https://free.law/2016/10/10/the-cost-of-pacer-data-around-one-billion-dollars/>

Enfin, rappelons que si un algorithme peut prédire une grande partie des dispositifs, c'est à l'évidence parce que les juges appliquent avant tout la loi qui est la même d'un cas à un autre.

Il faut donc rester vigilant sur les promesses de la justice prédictive. Il s'agit assurément d'un outil puissant qui nous apprendra beaucoup sur notre justice mais il faudra nous assurer que les différents biais possibles sont parfaitement connus et pris mieux en compte dans les résultats produits. Seule la transparence dans la démarche est à même d'atteindre ces objectifs.