

Quantification de l'aléa judiciaire, intelligence artificielle et applications

Mots-clés:

- 1 – apprentissage automatique
- 2 – jurisprudence
- 3 – modélisation

Par



Jacques Lévy Véhel,
Docteur en Mathématiques,
Directeur de Recherches INRIA,
Président de Case Law Analytics

L'intelligence artificielle permet de modéliser certains aspects de l'activité juridique. En particulier, en appliquant une méthodologie rigoureuse, on peut présenter l'éventail des décisions qui seraient prises par une juridiction donnée sur un dossier caractérisé par quelques dizaines de critères. Cette quantification de l'aléa judiciaire s'ajoute à la panoplie des outils à la disposition des professionnels du droit et leur permet d'analyser finement les stratégies des parties et les pratiques des magistrats.

C'est devenu un lieu commun : comme quasiment la totalité des autres activités humaines, le droit est à son tour en train d'être bouleversé par les progrès récents d'un ensemble de techniques généralement désignées sous le nom *d'intelligence artificielle*. Il est crucial que les juristes, dont la culture est, contrairement par exemple aux médecins, fort différente de celle des mathématiciens, appréhendent clairement les apports et limites des nouveaux outils qui s'invitent dans leur champ, afin de procéder à une analyse pertinente et sereine. A cet effet, une compréhension minimale de leur fonctionnement est un préalable afin que le dialogue nécessaire avec la communauté scientifique s'engage sur des bases propres à assurer que l'irruption de ces technologies se traduise le plus rapidement possible par un meilleur fonctionnement du système judiciaire, et non par une dégradation de celui-ci.

I. La quantification de l'aléa judiciaire est possible dans certaines limites.

A. Il est souhaitable de donner à voir les pratiques des juridictions pour éviter des attentes irréalistes des parties.

Si on se plonge dans la jurisprudence dans divers domaines, on constate rapidement que les parties arrivent régulièrement au procès avec des prétentions exagérément éloignées. Il n'est pas rare de voir par exemple des dossiers où la victime d'un dommage corporel demande plusieurs centaines de milliers d'euros au titre des souffrances endurées alors que l'assureur n'en propose quelques dizaines de milliers. Les mêmes différences d'ordre de grandeur se

retrouvent couramment pour des prestations compensatoires. En droit commercial, on trouve de nombreux cas dans lesquels la victime d'une rupture brutale d'une relation commerciale établie demande plus de vingt mois de préavis quand l'auteur ne propose aucune compensation.

Plus globalement, si on regarde l'intégralité des arrêts d'appel concernant des prestations compensatoires depuis 2012, on constate que dans environ 10 % des cas, les demandes et propositions ne sont pas du même ordre de grandeur. Les mêmes chiffres valent sur la totalité des arrêts pris par la Cour d'Appel de Paris en ce qui concerne les ruptures brutales de relations commerciales établies.

Ces considérations incitent à penser que fournir des outils aux professionnels du droit pour les aider à trouver des points d'équilibre en présentant des anticipations raisonnables de ce que pourra décider une cour sur un dossier donné serait faire œuvre utile.

Calculer ces anticipations est ce que nous désignons dans la suite par l'expression « quantification de l'aléa judiciaire ».

B. Il faut se garder de faire un usage excessif de la quantification de l'aléa.

Proposer de tels outils implique un engagement éthique fort : il faut assurer la validité des anticipations présentées, faire en sorte que leur utilisation n'entraîne pas à terme un biais dans les pratiques des avocats comme des magistrats, et enfin accompagner les professionnels du droit pour les aider à les maîtriser au mieux de l'intérêt de tous.

Mais surtout, contre les phantasmes et les arguments marketing parfois avancés, il faut sans cesse rappeler l'essence d'une décision de justice, à savoir qu'au terme d'une procédure, un juge appliquera les règles de droit en les individualisant, après avoir entendu les avocats des deux parties. La quantification de l'aléa intervient bien en amont, comme une aide à la préparation d'un dossier pour un avocat, et en tant qu'indication sur les pratiques de sa juridiction pour un magistrat. En particulier, aucune anticipation ne devrait jamais limiter la créativité de l'avocat ni contraindre la décision du juge.

II. Les techniques mises en œuvre empruntent aux derniers développements de l'intelligence artificielle tout en révélant des fonctionnements parfois inattendus.

A. L'intelligence artificielle peut s'appliquer en droit, à condition de respecter une méthodologie scientifique rigoureuse et de tenir compte des spécificités de ce domaine.

La méthodologie appliquée pour quantifier l'aléa judiciaire est identique à celle employée dans d'autres champs, comme par exemple la médecine. Loin de se contenter, comme on le croit parfois, de simples statistiques plus ou moins fines sur les décisions passées, elle repose sur une modélisation du processus de prise de décision. A cette fin, elle se déploie en trois étapes :

- après avoir choisi un domaine bien circonscrit (par exemple la concurrence déloyale, les licenciements sans cause réelle et sérieuse ou les sanctions disciplinaires), la première phase, de pure expertise juridique, consiste à faire une liste des principaux critères, notamment textuels et jurisprudentiels, sur lesquels s'appuie le magistrat pour prendre sa décision. Les jugements comportent nombre de ces déterminants, mais pas

tous. La plupart sont évidents puisqu'ils reposent sur l'application de la norme, mais certains sont plus factuels. L'un des intérêts de l'approche décrite ici est d'ailleurs de rendre explicites ces déterminants. Si l'identification de ces critères relève exclusivement de la mission du juriste, et plus précisément des professionnels du contentieux concerné, c'est un exercice qui demande un certain état d'esprit et une grande rigueur, tant il est crucial de refléter de manière la plus neutre possible les caractéristiques les plus souvent mises en avant dans les décisions prises dans ce domaine particulier. Suivant les cas, le nombre de critères varie d'une trentaine à une centaine ;

- dans la seconde étape, on analyse une quantité significative de jurisprudence (typiquement de l'ordre de quelques centaines à quelques dizaines de milliers de décisions) à la lumière des critères précédemment définis. Il peut s'agir soit de jugements de première instance soit d'arrêts d'appel ;
- la modélisation proprement dite est l'objet de la troisième phase : on utilise l'intelligence artificielle, et plus précisément les méthodes les plus récentes d'apprentissage automatique, pour reproduire le plus fidèlement possible les décisions prises par les magistrats en combinant les critères choisis à la première étape. Il est crucial de noter que l'on n'enseigne nullement à la machine les règles de droit sur lesquelles est censée se fonder la décision : ceci impliquerait en effet de laisser un algorithme *dire le droit*, et ouvrirait la voie à un monde dans lequel on accepterait d'être jugés par des intelligences artificielles. Ce serait nier une fonction fondamentale de l'acte de juger, qui est de produire du symbolique et de réguler les rapports dans la société. Au contraire, suivant le principe du *réalisme juridique*, on se contente de lui apprendre à produire des décisions indistinguables de celles qui sont prises, via des raisonnements juridiques, par des humains.

Une spécificité du droit complique singulièrement la méthodologie générale décrite ci-dessus : il est en effet essentiel de tenir compte du fait que, pour un jeu de critères donnés, il n'y a pas une seule décision que serait la « bonne » solution. Cela tient à plusieurs facteurs, et nous nous contenterons d'en mentionner deux : en premier lieu, même avec un grand nombre de critères, on ne peut prétendre caractériser totalement un dossier. Il existe toujours des éléments spécifiques qui échappent à toute catégorisation. Ensuite, et plus profondément, même en imaginant deux cas strictement identiques, parce que la justice est humaine et que la décision rendue est le fruit du travail des avocats et de l'interprétation du juge à un instant donné, parce que les plaidoiries seront différentes, parce que le délibéré est le fruit d'une concertation et d'une discussion entre plusieurs magistrats, il est normal et souhaitable que les issues puissent différer si le même dossier est présenté deux fois.

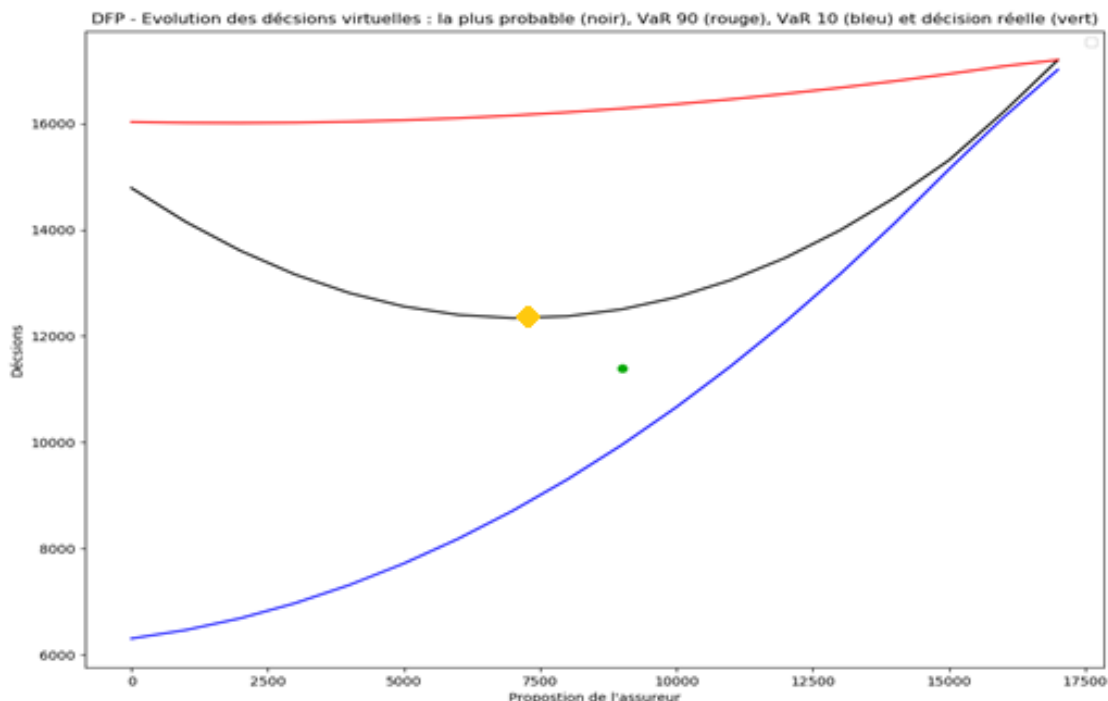
Pour tenir compte de ces aspects, l'apprentissage automatique doit être conçu de sorte à modéliser non pas une seule réponse, mais un éventail de jugements possibles, qui seront, *de façon prouvée*, ceux qui seraient rendus sur un dossier donné devant une juridiction donnée. Autrement dit, la tâche assignée à l'intelligence artificielle est de présenter toutes les décisions qui pourraient être prises dans les dossiers ayant les mêmes critères explicités dans la première phase. C'est dans ce sens précis que l'on parle de quantification de l'aléa : sans rien connaître

d'une affaire, on ne peut évidemment émettre le moindre diagnostic. A l'opposé, un avocat aguerri pourra souvent, après analyse approfondie d'un dossier et avec une connaissance fine de la jurisprudence, des pratiques de sa juridiction (y compris éventuellement la personnalité du juge ou de la chambre) et enfin des compétences du conseil de la partie adverse, prévoir assez précisément ses chances de succès. Entre ces deux extrêmes, la machine va présenter les diverses issues envisageables compte tenu des connaissances partielles dont elle dispose. Elle réduit le champ des possibles en utilisant les informations fournies par l'ensemble des critères, et permet ainsi un premier diagnostic rapide et fiable sur lequel le juriste peut s'appuyer pour approfondir sa réflexion.

B. La modélisation fait apparaître divers « motifs » juridiques ou humains qu'il est intéressant d'explorer.

La modélisation ci-dessus ouvre de nombreuses perspectives inédites, comme par exemple la possibilité d'analyser finement la manière dont l'éventail des décisions possibles varie quand on modifie un seul critère dans un dossier. On peut ainsi exposer diverses disparités, par exemple géographiques, et essayer de comprendre si elles sont légitimes ou au contraire devraient être corrigées. Ainsi, loin de perpétuer des biais, une critique souvent faite à l'encontre des systèmes fondés sur l'intelligence artificielle, cette approche permet au contraire de les mettre en évidence et d'agir pour les effacer le cas échéant.

Nous allons ici nous concentrer sur un autre phénomène que nous baptisons « sourire d'indemnisation », en référence à la forme de la courbe noire ci-dessous :



Pour rendre les choses concrètes, considérons un exemple : on entend parfois que, dans les dossiers de préjudice corporel, l'assureur a intérêt à faire une proposition « raisonnable » s'il veut limiter le montant d'indemnisation qui sera effectivement décidé par le juge. Autrement

dit, faire une offre trop basse peut l'exposer à devoir payer plus que s'il avait choisi de se montrer plus généreux.

La modélisation, qui, dans ce domaine, est assez fine puisqu'elle se fonde sur une centaine de critères, permet de tester la validité de cette hypothèse. Les graphes ci-dessus décrivent les montants virtuels alloués au titre du déficit fonctionnel permanent dans un dossier de dommage corporel. Les courbes sont obtenues de la façon suivante : on a repris les éléments d'un dossier effectivement soumis à la Cour d'Appel de Paris, sur lequel on a interrogé la machine en faisant varier un seul critère, à savoir le montant d'indemnisation proposé par l'assureur. Celui-ci peut en principe varier entre zéro et le montant demandé par la victime, ici 17 500 euros. Comme expliqué plus haut, l'intelligence artificielle va produire, pour chaque simulation, l'éventail des décisions qui seraient prises par la Cour. Le graphe noir représente l'évolution du montant le plus probable en fonction du montant proposé par l'assureur. Par exemple, le premier point de la courbe, à gauche, nous informe du fait que, si l'assureur ne compte pas indemniser la victime au titre de ce poste de préjudice (valeur 0 sur l'axe horizontal), le montant le plus probable décidé par la Cour sera d'environ 15 000 euros (valeur sur l'axe vertical). Quand, à l'opposé (le point noir le plus à droite de la courbe), l'assureur propose un montant de 17 500 euros, qui coïncide avec la demande de la victime, la Cour accorde évidemment ce même montant. Comme on le voit, le graphe noir prend peu ou prou la forme d'un sourire, ce qui signifie qu'il y a bien une proposition « optimale » (figurée par le symbole jaune sur le graphe) qui minimise le montant le plus probable que l'assureur devra payer, ici à peu près 12 500 euros s'il propose un peu moins de 7 500 euros (valeurs respectivement sur les axes verticaux et horizontaux).

Il est tout aussi important de remarquer que le sourire n'apparaît pas si l'on considère les décisions les plus et les moins « sévères » : la courbe bleue représente le montant maximal que l'assureur devra payer selon les 10 % de magistrats accordant l'indemnité la plus faible, et la rouge le montant maximal que l'assureur devra payer selon les 10 % de magistrats accordant l'indemnité la plus forte. Dans les deux cas, les montants croissent régulièrement, ce qui invalide l'hypothèse d'un effet sourire.

Ainsi, la réponse à la question « est-ce que l'assureur a intérêt à se montrer généreux pour minimiser le montant qu'il devra payer pour indemniser la victime » est : « cela dépend ». C'est oui si l'on se fonde sur le montant le plus probable, et non si l'on considère les valeurs extrêmes.

L'effet sourire est observé, avec des variantes, dans d'autres domaines, comme par exemple les ruptures brutales de relations commerciales établies, quand on fait varier le critère « durée du préavis proposé par l'auteur de la rupture ».

L'explication de ce phénomène dépasse non seulement le cadre de cet article mais aussi et surtout les compétences de l'auteur. On se contentera de noter que, lorsque l'on interroge les magistrats à ce sujet, on obtient trois catégories de réponses : une partie ne se reconnaît pas dans le sourire d'indemnisation et semble peu sensible dans sa prise de décision à la proposition de l'assureur. Une seconde admet qu'un certain agacement devant une mauvaise foi évidente peut conduire à un tel effet. Enfin, la troisième explique qu'au moment de prendre la décision, compte est tenu des positions des deux parties, qui servent de points de référence approximatifs. Si l'une de ces positions est manifestement déraisonnable, la référence disparaît et il est naturel que la décision s'en ressente.

Il est intéressant de remarquer que ces trois réactions sont parfaitement compatibles avec les courbes présentées ci-dessus : on peut imaginer que les magistrats de la première catégorie font partie de ceux qui décident de montants « extrêmes », alors que les autres se prononcent pour des indemnisations plus fréquemment observées.

L'étude qui vient d'être faite est évidemment trop frustrante pour qu'on puisse en tirer des conclusions générales. Elle illustre cependant le potentiel d'une modélisation mathématique fine et l'enrichissement que celle-ci pourrait apporter si des juristes s'en saisissent pour explorer les pratiques des divers acteurs du droit et mettre en évidence des effets peut-être intéressants.