

PREDICTING EMPLOYMENT NOTICE PERIOD WITH MACHINE LEARNING: PROMISES AND LIMITATIONS

*Samuel Dahan, Jonathan Touboul,
Jason Lam, and Dan Sfedj**

Rapid advances in data analysis techniques—particularly for predictive algorithms—have opened the door for radically new perspectives on legal practice and access to justice. Several firms in North America, Asia, and Europe have set out to use machine-learning techniques to generate legal predictions, raising concerns regarding ethics, reliability and limits on prediction accuracy, and potential impact on case law development. To explore these opportunities and challenges, we consider in depth one of the most litigated issues in Canada: wrongful termination disputes and, more specifically, the question of reasonable notice determination. Beyond the thorough analysis of this question, this paper is also intended to act as a road map for non-technicians (and especially lawyers) on the application of artificial intelligence (AI) methods, illustrating both their potential benefits and limitations in other areas of dispute resolution.

To achieve these results, we first created a large dataset by annotating historic cases related to employment termination. This dataset proved useful for assessing the *predictability* of reasonable notice of termination, that is, the accuracy and precision of AI predictions. In particular, it helped identify the degree of inconsistency in notice period cases, incidentally exposing the limitations of legal predictions. We then developed predictive algorithms to estimate notice periods based on details of the employment period and investigated their accuracy and performance. Moreover, we thoroughly analyzed these algorithms to better understand the judicial process, and in particular to quantify the weight and influence of case-specific features in the determination of reasonable notice. Finally, we closely analyzed cases that were poorly predicted by the algorithms to understand the judicial decision-making process and identify inconsistencies—a strategy that will ultimately yield a deeper practical understanding of case law.

This project will open the door to the development of an access-to-justice project and will provide users with an open-access platform for employment legal help (www.MyOpenCourt.org).

Les progrès rapides des techniques d'analyse de données — les algorithmes prédictifs en particulier — ont ouvert la porte à des avenues radicalement nouvelles en matière de pratiques juridiques et d'accès à la justice. Plusieurs cabinets d'avocats d'Amérique du Nord, d'Asie et d'Europe ont entrepris d'utiliser des techniques d'apprentissage statistique (*machine learning*) pour prédire et générer des conclusions d'ordre juridique, ce qui soulève des préoccupations concernant l'éthique, la fiabilité et les limites de la précision de ces conclusions, ainsi que leur impact potentiel sur le développement de la jurisprudence. Pour explorer ces possibilités et ces défis, nous examinons en profondeur l'une des questions les plus litigieuses au Canada : les licenciements abusifs et, plus particulièrement, la question de la détermination du préavis raisonnable. Au-delà de l'analyse approfondie de cette question, cet article se veut également une feuille de route pour les non-techniciens (et surtout les avocats) sur l'application des méthodes d'intelligence artificielle (IA), illustrant à la fois leurs avantages potentiels et leurs limites dans d'autres domaines de la résolution des litiges.

Pour atteindre cette fin, nous avons d'abord colligé un vaste ensemble de données en annotant les cas historiques de congédiement injustifiés. Cet ensemble de données s'est avéré utile pour évaluer la prévisibilité d'un préavis raisonnable de licenciement, c'est-à-dire la précision des prédictions de l'IA. En particulier, cette approche permet de déterminer le degré d'incohérence et de variation des cas de préavis, en exposant incidemment les limites de ses conclusions légales. Nous avons développé des algorithmes prédictifs afin d'estimer les délais de préavis en fonction de la durée de l'emploi et avons étudié leur précision et leur performance. De plus, nous avons procédé à une analyse approfondie de ces algorithmes afin de mieux comprendre le processus judiciaire, et en particulier de quantifier le poids et l'influence des caractéristiques propres à chaque affaire dans la détermination du préavis raisonnable. Enfin, nous avons analysé de près les cas mal prédits par les algorithmes d'IA afin de mieux comprendre le processus décisionnel judiciaire et d'en déterminer les incohérences — une stratégie qui permettra en définitive d'approfondir la compréhension pratique de la jurisprudence.

Ce projet ouvre la voie au développement d'un projet d'accès à la justice à plus grande échelle et fournira aux utilisateurs une plateforme en libre accès d'aide juridique en droit du travail (www.MyOpenCourt.org).

* Samuel Dahan, Conflict Analytics, Queen's Law Faculty, Cornell Law School; Jonathan Touboul, Conflict Analytics, Brandeis University, Department of Mathematics and Volen National Center for Complex Systems; Jason Lam, Queen's School of Computing; Dan Sfedj, CentraleSupélec Paris and Brandeis University, Department of Mathematics. [continued on next page]