

LE ROLE FUTUR DES MEGADONNEES ET DE L'APPRENTISSAGE AUTOMATIQUE DANS L'AMELIORATION DE L'EFFICACITE DES INSPECTIONS

Introduction

Les inspections sont probablement l'instrument politique le plus important qu'utilisent les inspections du travail gouvernementales afin de s'assurer que les entreprises prennent les mesures nécessaires pour respecter la législation en matière de sécurité et de santé au travail. Toutefois, l'effet qu'ont les inspections dépend de plusieurs facteurs. L'un des facteurs fondamentaux est le processus de sélection des objets à inspecter, c'est-à-dire les entreprises ou les sites à inspecter. En principe, il existe au moins trois approches de sélection différentes. Une première approche consiste à inspecter toutes les entreprises quels que soient le risque potentiel, la taille de l'entreprise, le type d'industrie ou tout autre critère. Une deuxième approche consiste à sélectionner les entreprises sur la base d'un échantillonnage aléatoire, chaque entreprise, indépendamment de toute caractéristique, ayant une probabilité égale d'être sélectionnée. En ce qui concerne les conditions préventives et économiques, ces deux méthodes sont généralement considérées comme inefficaces (Blanc, 2013). Par conséquent, la plupart des inspections du travail sélectionnent les objets sur la base d'une troisième approche, à savoir l'approche fondée sur les risques. En bref, l'approche fondée sur les risques implique la sélection des objets à inspecter sur la base du niveau de risques.

Bien que l'approche fondée sur les risques soit un principe essentiel pour la plupart des inspections du travail modernes, son application dans la pratique pose des défis de taille. La raison principale réside dans l'absence de méthodes d'analyse des risques suffisamment fines (Mischke et al., 2013). En l'absence de méthodes appropriées pour rendre possible l'établissement de priorités selon les risques, l'approche fondée sur les risques est susceptible de devenir une déclaration de politique gouvernementale sans conséquences pratiques tangibles. Il est donc nécessaire de développer des méthodes qui permettent de cibler les entreprises à hauts risques (Weil, 2008).

La plupart des inspections du travail collectent et stockent d'énormes quantités de données sur les objets à inspecter et leurs activités d'inspection. Ainsi, les services d'inspection possèdent potentiellement des volumes de données importants et en augmentation rapide, que l'on appelle aujourd'hui les « mégadonnées ». Les mégadonnées, combinées à la technologie de l'apprentissage automatique, sont de plus en plus utilisées à différentes fins prédictives, en retirant un enseignement des tendances masquées au sein des données. Par exemple, la valeur prédictive des mégadonnées et des techniques d'apprentissage automatique est testée dans des domaines aussi divers que le pronostic du cancer et les résultats pour les patients, la prévision des faillites, la prévision des prix du pétrole, la détection de la fraude fiscale, la prévision de la criminalité et les prévisions boursières. La question fondamentale abordée dans le présent document est toutefois de savoir si l'utilisation des mégadonnées et de la technologie de l'apprentissage automatique pour cibler les objets à inspecter présentant des risques élevés constitue une voie prometteuse pour les inspections du travail.

Ciblage fondé sur les risques

Selon les principes de bonnes pratiques de l'Organisation de coopération et de développement économiques pour la politique de la réglementation (OCDE, 2014), l'analyse et l'évaluation des risques devraient servir de base au ciblage des objets à inspecter par les inspections du travail. Cela signifie que les entreprises devraient être sélectionnées en vue d'une inspection sur la base

des évaluations de la probabilité et des conséquences de facteurs de risque tels que les accidents, l'exposition à des conditions nocives et les conditions de travail illégales. Le ciblage fondé sur les risques repose essentiellement sur la reconnaissance du fait qu'il n'est pas possible de contrôler tous les domaines à risques et tous les objets à risques en raison de ressources d'inspection limitées. Pour ce qui est des inspections de sécurité et de santé des autorités de l'inspection du travail, cela signifie que certains domaines problématiques doivent être prioritaires par rapport à d'autres. En outre, certaines entreprises doivent être inspectées en priorité par rapport à d'autres.

Le principe du ciblage fondé sur les risques n'est pas nouveau. Il y a près de 50 ans, lors de l'évaluation par le Comité Robens du système britannique de surveillance de la sécurité et de la santé au travail, l'approche fondée sur les risques (combinée à l'autoréglementation) a été introduite en tant qu'idéal dans le processus de modernisation de l'inspection réglementaire (Robens, 1972). Afin d'assurer l'utilisation rentable des ressources d'inspection, le rapport Robens recommandait à l'autorité réglementaire de concentrer de manière sélective ses ressources sur les domaines problématiques les plus graves et de donner la priorité aux entreprises et aux problèmes qui avaient été identifiés par l'analyse systématique de toutes les données disponibles en matière de sécurité et de santé, par exemple les statistiques concernant les accidents, les informations techniques et les connaissances locales des services d'inspection.

Les recommandations du rapport Robens ont été largement adoptées par les inspections du travail au niveau international et dans les États membres de l'UE. La généralisation de l'approche fondée sur les risques signifie que la plupart des inspections du travail modernes ont adopté l'idée de ne plus affecter de ressources aux objets à faibles risques et de concentrer davantage de ressources aux objets présentant les risques les plus élevés. Pour que cela soit possible, un certain type d'analyse des données est nécessaire. Les méthodes d'analyse permettant d'identifier les industries à hauts risques et les groupes de travailleurs exposés à des risques sont bien développées. Ces analyses fondées sur les risques s'appuient généralement sur des statistiques nationales relatives, par exemple, aux maladies professionnelles, aux accidents du travail et aux expositions professionnelles. Les analyses constituent le fondement des campagnes d'inspection, des plans stratégiques et des domaines prioritaires nationaux, voire internationaux.

Les méthodes qui permettent d'établir des priorités entre les entreprises d'un secteur d'activité sont beaucoup moins courantes que les analyses fondées sur les risques à caractère général. Parmi les inspections du travail, une approche commune pour cibler les entreprises exposées à des risques concrets consiste à s'appuyer sur les connaissances locales des inspecteurs. Certaines inspections du travail, comme celles du Danemark et de Suède, ont étudié la possibilité d'utiliser des systèmes de classement des risques sur la base d'échelles additives. En utilisant des échelles additives, chaque entreprise reçoit des scores de risques basés sur plusieurs caractéristiques de l'entreprise (par exemple, la taille, le type d'industrie et le nombre d'accidents enregistrés) qui sont additionnés pour former un score global, et les entreprises ayant le score global le plus élevé sont inspectées en priorité. Toutefois, le problème que pose l'utilisation de telles échelles additives est qu'elles présentent des niveaux relativement faibles de validité prédictive, c'est-à-dire que le score n'est pas particulièrement approprié pour distinguer les entreprises à hauts risques de celles à faibles risques.

Mégadonnées et apprentissage automatique

Le processus d'établissement des priorités entre les entreprises équivaut à chercher une aiguille dans une botte de foin. Dans ce cas, la botte de foin est composée, potentiellement, de centaines de milliers d'objets à inspecter possibles, mais seul un certain nombre de ces objets sont des aiguilles, c'est-à-dire qu'ils présentent un niveau de risque intolérable. Trouver des aiguilles dans une botte de foin, c'est en grande partie ce que sont les mégadonnées et l'apprentissage automatique.

L'objectif principal des algorithmes d'apprentissage automatique est de fournir un modèle statistique qui puisse être utilisé pour effectuer des prédictions, des classifications, des estimations ou des tâches similaires. Par exemple, dans le domaine de la prédiction du cancer, les chercheurs utilisent depuis plus de trois décennies des algorithmes d'apprentissage automatique pour prédire la prédisposition au cancer, le taux de récurrence du cancer et le taux de survie au cancer. D'un point de vue thématique, la prédiction du cancer est très éloignée du ciblage fondé sur les risques des objets à inspecter. Cependant, ils sont tous deux des exemples de défis prédictifs ou de problèmes du type «chercher une aiguille dans une botte de foin».

Les deux principaux types communs d'algorithmes d'apprentissage automatique sont l'apprentissage supervisé et l'apprentissage non supervisé. Dans l'apprentissage supervisé, l'algorithme consiste en une variable dépendante (par exemple, le niveau de risque) qui doit être prédite à partir d'un ensemble de variables indépendantes. Bien entendu, des prédictions précises exigent des niveaux élevés de corrélation entre les variables indépendantes et la variable dépendante. Dans l'apprentissage non supervisé, il n'y a pas de variable dépendante à prédire, mais l'objectif de l'algorithme est de classer les données en groupes (par exemple, différents groupes de risques) par similitude. Contrairement aux échelles additives, telles que celles explorées par les inspections du travail du Danemark et de Suède, les algorithmes utilisés dans l'apprentissage automatique améliorent progressivement leurs prévisions, principalement par tâtonnements. Cela signifie que l'apprentissage automatique se fait à partir des succès (prédictions correctes) et des erreurs (prédictions erronées) obtenus antérieurement et tente d'enregistrer ces connaissances pour rendre les prédictions plus précises en fonction de la rétroaction.

Utilisation des mégadonnées et de l'apprentissage automatique dans la sélection des objets à inspecter

Les algorithmes d'apprentissage supervisé et non supervisé nécessitent un volume suffisant de données, tant en ce qui concerne le nombre d'observations que le nombre de variables, généralement appelées «caractéristiques». Comme nous l'avons déjà indiqué, la plupart des inspections du travail collectent et stockent d'énormes quantités de données sur les objets à inspecter et leurs activités d'inspection. Les données disponibles concernent généralement des caractéristiques propres à l'entreprise, telles que le nombre d'employés, l'âge de l'entreprise, le groupe industriel, le nombre d'inspections précédentes, les résultats des inspections précédentes et les notifications d'accidents. De plus, la quantité de données augmente de jour en jour à mesure que les résultats de nouvelles inspections sont ajoutés. En principe, la résolution du problème de ciblage des entreprises à hauts risques en utilisant des mégadonnées devrait donc, au moins à première vue, être bien adapté aux algorithmes d'apprentissage automatique. Malgré cela, il n'y a eu que peu de tentatives de ce genre. Il existe toutefois quelques exceptions notables, qui illustrent toutes que les mégadonnées et l'apprentissage automatique pourraient être très utiles aux inspections du travail pour résoudre le problème du ciblage des objets à hauts risques à inspecter.

Le premier exemple est une étude de recherche qui s'est penchée sur la pertinence des méthodes d'apprentissage automatique pour la prévision des accidents du travail ou, plus précisément, des chutes au sol (Matías et al., 2008). Malgré ses prédictions relativement précises, l'inconvénient de cette étude est que les caractéristiques incluses dans les algorithmes ne correspondent pas au type de données dont les inspections du travail disposent normalement (par exemple, l'utilisation d'équipements de protection individuelle et les pratiques d'entretien de la propreté). En outre, les chutes au sol ne représentent qu'une infime partie des risques professionnels concernés par les inspections du travail.

Le deuxième exemple est également une étude issue de la recherche (Hajakbari et Minaei-Bidgoli, 2014). Cette étude a développé un système de notation pour prédire le risque d'accidents du travail. En outre, l'étude a conclu qu'il est possible de prédire le risque de différents types

d'accidents du travail de manière relativement précise, sur la base de certaines caractéristiques générales de l'entreprise (activité principale de l'entreprise, répartition hommes-femmes, nombre de salariés, etc.). En outre, l'étude a conclu que l'utilisation d'un algorithme permettrait d'identifier les lieux de travail qui nécessitent des inspections de sécurité et de santé de périodiques. Les données utilisées dans cette étude ont été extraites de la base de données d'un service d'inspection du travail. L'inconvénient de l'étude, cependant, est encore une fois que les accidents du travail ne représentent qu'un des nombreux risques sur le lieu de travail que traitent les inspections du travail. En outre, le recours à des statistiques relatives aux blessures pose un problème particulier, en ce sens que ces données sont connues pour être très vulnérables à la sous-déclaration.

Le troisième exemple est un outil développé par l'Autorité de l'inspection du travail norvégienne (NLIA) pour aider les inspecteurs à sélectionner les entreprises en fonction des risques (Dahl et al., 2018). L'outil, appelé «outil de prédiction des catégories de risque» (Risk Group Prediction Tool, RGPT), différencie les entreprises en quatre catégories en fonction des risques prévisibles: les entreprises présentant les risques les plus faibles, des risques faibles, des risques élevés et les risques les plus élevés. Plus la catégorie de risque d'une entreprise donnée est élevée, plus il est probable qu'une inspection future de cette entreprise permettra d'identifier de sérieux manquements au respect de la législation en matière de sécurité et de santé. La catégorie qui est assignée à une entreprise est accessible aux inspecteurs par le biais de l'interface utilisateur en ligne interne de la NLIA. Ainsi, lorsqu'ils ciblent les entreprises à inspecter, les inspecteurs sont informés des catégories de risque des entreprises et sont donc en mesure de faire des sélections en fonction des risques.

L'outil RGPT a été conçu sur la base d'une modélisation prédictive au moyen d'un algorithme d'apprentissage automatique utilisant une analyse dite de régression logistique binaire. Sur la base du modèle de régression, toutes les entreprises norvégiennes (environ 230 000) sont classées dans l'une des quatre catégories de risque. Cette classification se fait en deux étapes. Dans un premier temps, le modèle de régression prédit la probabilité qu'une inspection future permette d'identifier de sérieux manquements au respect de la législation en matière de sécurité et de santé. Lors de la deuxième étape, le modèle utilise la valeur de cette probabilité prédite pour attribuer à l'entreprise une catégorie de risque.

Au départ, l'outil a été développé à partir des enregistrements d'environ 35 000 inspections de sécurité et de santé effectuées par la NLIA. Cependant, les prédictions faites par cet outil deviennent progressivement et automatiquement plus précises à mesure que le nombre d'inspections augmente. Cela signifie que l'algorithme s'ajuste lui-même en fonction de la rétroaction (prévisions correctes ou erronées) qu'il reçoit lorsque de nouvelles inspections sont effectuées et enregistrées dans la base de données de la NLIA.

L'outil RGPT appartient à la classe des algorithmes d'apprentissage supervisé, où les inspections de sécurité et de santé constatant de sérieux manquements (variable dépendante) doivent être prédites à partir d'un ensemble de caractéristiques de l'entreprise. Les caractéristiques que l'outil RGPT utilise sont des caractéristiques générales de l'entreprise telles que la taille de l'entreprise, le groupe industriel, le nombre d'inspections précédentes, les résultats des inspections précédentes, l'âge de l'entreprise, la localisation géographique et les notifications d'accidents. La validité prédictive de l'outil est vérifiée tous les mois et l'expérience acquise jusqu'à présent (après environ 18 mois de tests) montre que l'algorithme parvient à cibler les entreprises à hauts risques avec une précision extrême. Cela signifie qu'il y a peu de «faux positifs» et peu de «faux négatifs», c'est-à-dire que peu d'inspections dans la catégorie présentant les risques les plus faibles donnent lieu à l'identification de sérieux manquements, alors que la grande majorité des inspections dans la catégorie présentant les risques les plus élevés donnent lieu à l'identification de sérieux manquements. Les groupes à faibles risques et à risques élevés se situent entre les deux extrêmes.

Les résultats de l'utilisation de l'outil développé par la NLIA montrent qu'il est possible de cibler les objets à inspecter en utilisant les mégadonnées et l'apprentissage automatique. Des approches similaires d'apprentissage automatique ont également été testées par au moins deux autres inspections du travail européennes avec des résultats prometteurs: l'Autorité suédoise de l'environnement de travail (Ridemar, 2018) et l'inspection néerlandaise SZW (Jacobusse et Veenman, 2016). Toutefois, l'outil norvégien n'est pas nécessairement transférable à d'autres inspections du travail, car son utilisation dépend de la manière dont les données sont stockées, de la qualité des données, de l'accès aux données et de la structure de la base de données. Par ailleurs, le ciblage des entreprises en utilisant cet outil implique l'acceptation de la manière dont les risques sont définis et opérationnalisés dans l'algorithme. Tel que décrit, l'outil est basé sur une définition des risques qui implique que plus la catégorie de risque d'une entreprise donnée est élevée, plus il est probable qu'une inspection de cette entreprise permettra d'identifier de sérieux manquements au respect de la législation en matière de sécurité et de santé. Cela signifie que l'outil concerne principalement les risques dits de gestion et de contrôle et non les risques inhérents. Alors que les risques de gestion et de contrôle découlent de la capacité et de la volonté d'une entreprise de gérer les risques (par exemple, en respectant la législation pertinente), les risques inhérents sont ceux qui découlent de la nature des activités d'une entreprise (par exemple, chute de hauteur, exposition chimique et fatigue musculosquelettique).

En pratique, les risques de gestion et de contrôle et les risques inhérents sont liés. Toutefois, cela n'implique pas que les deux types de risques soient nécessairement fortement corrélés de manière empirique. Par conséquent, s'appuyer aveuglément sur des outils qui ciblent les entreprises en fonction d'un type de risques peut avoir pour conséquence le fait qu'un autre type n'est pas pris en compte. Dans le cadre du régime réglementaire norvégien, ce problème est contourné en privilégiant les risques inhérents lors de l'identification des domaines prioritaires, des groupes de travailleurs exposés à des risques et des industries à hauts risques, tandis que les risques de gestion et de contrôle sont privilégiés lors du ciblage spécifique des entreprises.

Défis à relever

Le fait que les risques de gestion et de contrôle, d'une part, et les risques inhérents, d'autre part, ne sont pas nécessairement corrélés de manière empirique, nous amène à relever un autre défi, probablement encore plus grand, dans l'application des algorithmes de traitement des mégadonnées et d'apprentissage automatique au ciblage fondé sur les risques. Les trois exemples d'outils d'apprentissage automatique ci-dessus sont des exemples de ciblage unidimensionnel, c'est-à-dire un ciblage fondé sur une définition particulière et une opérationnalisation des risques. Les risques dans le monde du travail ne sont cependant pas d'un seul type. Par conséquent, les autorités chargées de faire respecter la loi sont concernées par de multiples types de risques, tels que les accidents, l'exposition chimique, l'exposition biologique, les menaces psychosociales, les facteurs de risques musculosquelettiques et le dumping social. Parmi ces types de risques, il y a encore d'autres sous-types. L'élaboration de modèles de risques qui permettent de rendre compte de cette diversité est très difficile, car les différents types de risques ne sont pas nécessairement corrélés. Par conséquent, la représentation de cette diversité est très différente de la prévision de la probabilité d'un type particulier de risques (Dahl et al., 2018).

Un deuxième défi, mais connexe, rend encore plus complexe la tâche du ciblage fondé sur les risques. C'est ce qu'on appelle l'écueil politique (Black, 2010). Même si les algorithmes d'apprentissage automatique sont dynamiques, en ce sens qu'ils peuvent apprendre et s'adapter de leurs succès et de leurs erreurs à s'y adapter, ils ne peuvent pas prendre en considération des points de vue politiques différents. Premièrement, le contexte politique est instable. Ainsi, les types de risques qui méritent d'être classés comme prioritaires aujourd'hui pourraient ne pas l'être demain. Deuxièmement, le contexte politique possède de multiples facettes. Ainsi, les différentes parties prenantes, par exemple les politiciens, les employeurs, les employés, les médias et le grand public, ont des points de vue différents sur les types de risques qu'il convient de classer

comme étant prioritaires. Cela montre que les risques du monde du travail ne sont pas nécessairement une entité objective, mais une construction sociale.

Un troisième défi, qui mérite d'être pris en considération, est lié au fait que, même si les inspections du travail disposent d'énormes quantités de données sur les objets à inspecter, ces données concernent généralement des données recueillies au niveau de l'entreprise qui ne sont pas nécessairement les données les plus appropriées à considérer (voir, par exemple, Gunningham et Sinclair, 2007). Dans une base de données, une entreprise unique est généralement identifiée par un identifiant unique tel qu'un numéro d'organisation. La capacité d'un algorithme d'apprentissage automatique à attribuer à une entreprise donnée une valeur donnée relative aux risques prévisibles dépend d'identifiants uniques. Toutefois, tous les objets à inspecter potentiels ne sont pas automatiquement identifiables au moyen d'un identifiant unique. Par exemple, dans l'industrie de la construction, ce n'est pas nécessairement une entreprise de production de béton qui est visée par l'inspection, mais un chantier de construction temporaire. Il existe au moins deux défis liés à ce caractère temporaire. Premièrement, les chantiers de construction et autres lieux de travail temporaires pourraient ne pas être identifiables au moyen d'identifiants uniques. Deuxièmement, même s'ils étaient identifiables, le caractère temporaire signifie qu'un algorithme d'apprentissage automatique pourrait ne pas avoir la possibilité d'apprendre de ses erreurs et de ses succès prédictifs avant que le chantier de construction ne cesse d'exister et que les entreprises qui intervenaient sur le chantier n'aient rejoint un autre chantier.

En conclusion

Les défis décrits ci-dessus montrent qu'il est très difficile de cibler les objets à inspecter à hauts risques en utilisant des mégadonnées et des techniques d'apprentissage automatique. Cependant, ces défis n'enlèvent rien à l'utilité de ces techniques dans le cadre d'une approche fondée sur les risques. Au contraire, les défis montrent que le ciblage fondé sur les risques ne pourra probablement pas s'appuyer entièrement sur des algorithmes d'apprentissage automatique. L'exemple norvégien ci-dessus l'illustre bien. Plutôt que de permettre à l'algorithme de choisir directement les objets, les inspecteurs sont autorisés à prendre des décisions tenant compte des risques sur la base des prédictions de l'algorithme. Cela implique la combinaison de l'intelligence artificielle et de l'intelligence humaine, où chacune vient compléter les points forts de l'autre. Lorsqu'il s'agit de prédire des événements sociaux complexes en général, la combinaison des deux types d'intelligence est probablement une nécessité.

Références

- Black, J., 2010. Risk-based regulation: Choices, practices and lessons being learnt. OCDE 2010 (éd.), Risk and regulatory policy: Improving the governance of risk. OCDE, Paris.
- Blanc, F., 2013. Inspection reforms: why, how, and with what results. OCDE, Paris.
- Dahl, Ø., Kilskar, S. S., Skarholt, K., Rosness, R., 2018. Risikobasert tilsyn i de nordiske arbeidstilsynene. [Inspections du travail basées sur les risques dans les pays nordiques.] Conseil nordique des ministres, Copenhague.
- Gunningham, N., Sinclair, D., 2007. Factors impinging on the effectiveness of the Mines Inspectorate. The Australian National University (Université nationale d'Australie), Canberra.
- Hajakbari, M. S., Minaei-Bidgoli, B., 2014. A new scoring system for assessing the risk of occupational accidents: A case study using data mining techniques with Iran's Ministry of Labor data. Journal of Loss Prevention in the Process Industries 32, pp. 443-453.
- Jacobusse, G., Veenman, C., 2016. On selection bias with imbalanced classes. International Conference on Discovery Science, Bari, pp. 325-340.
- Matías, J. M., Rivas, T., Martín, J. E., Taboada, J., 2008. A machine learning methodology for the analysis of workplace accidents. International Journal of Computer Mathematics 85, pp. 559-578.
- Mischke, C., Verbeek, J. H., Job, J., Morata, T. C., Alvesalo-Kuusi, A., Neuvonen, K., Clarke, S., Pedlow, R. I., 2013. Occupational safety and health enforcement tools for preventing occupational diseases and injuries. Cochrane Database of Systematic Reviews 2013, Issue 8, Art. No CD010183.
- OCDE, 2014. Principes de bonnes pratiques de l'OCDE pour la politique de la réglementation. Organisation de coopération et de développement économiques, Paris
- Ridemar, A., 2018. Decision support for SWEA inspections. Institut royal de technologie (KTH), Stockholm.
- Robens, L., 1972. Safety and Health at Work: Report of the Committee 1970-1972. Service d'édition des publications officielles du Royaume-Uni (Her Majesty's Stationery Office), Londres.
- Weil, D., 2008. A strategic approach to labour inspection. International Labour Review 147, pp. 349-375.

Auteurs: Øyvind Dahl, chercheur senior, SINTEF Digital, Département Sécurité et fiabilité, Trondheim, Norvège

Annick Starren, Agence européenne pour la sécurité et la santé au travail (EU-OSHA)