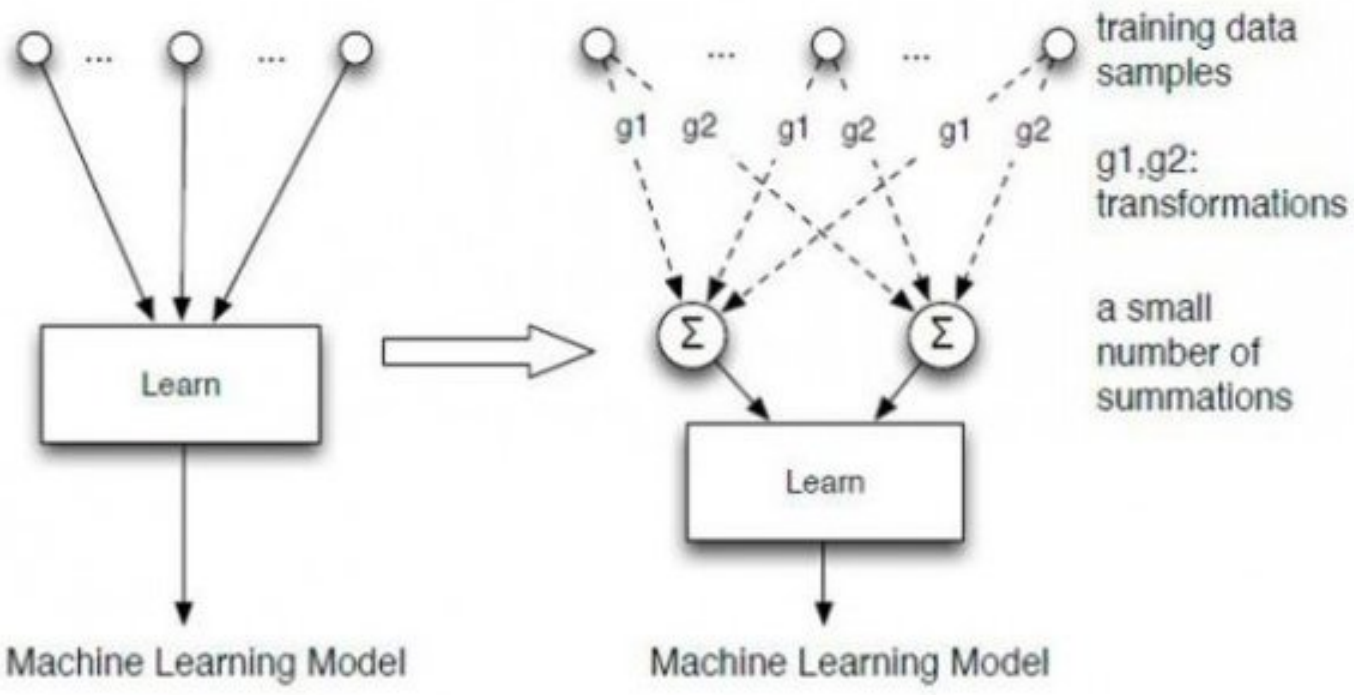
الكشف عن خوارزمية غير مألوفة في مجال الذكاء الصنعي

المصدر: جامعة ليهاي Lehigh University.

إنّ أنظمة تعليم الآلة أصبحت منتشرة في كل مكان. فهي الآن تتنبّأ بأحوال الطقس، وتتوقّع حدوث الزلازل، حتّى إنها تُطبِق على مكابح سيارتنا إذا لزم الأمر.

وللقيام بذلك، تقوم البنية البرمجيّة في هذه الأنظمة بحساب العلاقات التنبؤية التي تحتاجها من كميات هائلة من المعلومات، وتتعرّف الأنظمة على هذه العلاقات باستخدام خوارزميّات متطوّرة ـ مجموعة قواعد لحل مشاكل رياضيّة ـ بالإضافة إلى "بيانات التدريب **training data**". تستخدم هذه البيانات لاحقاً لبناء النماذج والخواص التي تسمح للنظام بتحديد أي الكتب الأكثر مبيعاً تفضّل قراءتها،

أو للتنبؤ باحتمالية هطول الأمطار في الأسبوع القادم.



تسمح الطريقة التي طورها كل من كاو Cao ويانغ Yang للأنظمة المتعلمة بأن "تنسى"، أو أن تحذف البيانات عبر إعادة حساب عدد ضئيل من عمليات الجمع بدلاً من إعادة بناء نماذج تتنبأ بالعلاقات بين البيانات الفردية. حقوق الصورة: ينجي كاو Yinzhi Cao وجونفن يانغ Junfeng Yang.

وغالباً، في هكذا عمليات معقدة سوف تمر البيانات الأولية عبر عمليات حسابية عديدة ضمن النظام. هذه العمليّات الحسابية بالإضافة إلى المعلومات المستمدة من البيانات تشكلان مع بعضهما شبكة انتشار معقّد **complex propagation network** تُدعى بـ"سلالة البيانات **Data's lineage**". إن أول من قام بصياغة هذا المصطلح هو ينجي كاو **Yinzhi Cao**، وهو أستاذ مساعد في الهندسة وعلوم الحاسب، بالإضافة لزميله جونفن يانغ **Junfeng Yang** من جامعة كولومبيا **Columbia University**، وهما يعملان على تطوير نهجٍ جديد لجعل أنظمة التعلّم تنسى.

نظراً لأهميّة هذا المفهوم في زيادة الأمان وحماية الخصوصيّة، يعتقد كل من كاو ويانغ أنّ التبني البسيط للأنظمة القادرة على النسيان سيكون مطلوباً بشكل متزايد. وقد قام هذان الباحثان بتطوير طريقة للقيام بذلك بسرعة وفعالية أكبر من الطرق المستخدمة حالياً.

حاز كل من هذين الباحثين على منحة المؤسّسة الوطنيّة للعلوم، والتي تقدر بـ 1.2 مليون دولار وذلك لتطويرهم مفهوم "إلغاء تعليم الآلة **machine unlearning**" وهو مفهوم واعد جداً.

يقول كاو، وهو الباحث الرئيسي في المشروع: "يجب على نظم النسيان الفعالة أن تكون قادرة على السماح للمستخدمين بتحديد بيانات معينة لنسيانها مع وجود مستويات مختلفة من التفصيل، ويتوجّب على هذه الأنظمة أيضاً حذف البيانات وإلغاء تأثيرها، وبالتالي ستبنى كل العمليّات المستقبليّة وكأنّ تلك البيانات لم تكن موجودة أبداً".

زيادة الأمان وحماية الخصوصيّة:

هناك أسباب عدة قد يتطلب فيها المستخدم العادي أو مزود الخدمة من النظام أن ينسى بيانات معينة بالإضافة لأصل هذه البيانات بالكامل. أحد هذه الأسباب هي الخصوصية **Privacy**.

حيث إنه بعد قيام موقع فيس بوك بتغيير سياسة الخصوصيّة خاصّته، قام العديد من المستخدمين بحذف حساباتهم والبيانات المتعلقة بها. بالإضافة إلى حادثة اختراق صور **iCloud** في العام 2014 –والتي أدت إلى تسريب الكثير من صور المشاهير الخاصة عبر مجموعة خدمات آبل **Apple** السحابيّة– والتي أدّت أيضاً إلى انتشار العديد من المقالات على الإنترنت لتعليم المستخدمين كيفيّة حذف الصور من على نظام **IOS** بشكل كامل، ومن ضمنها النسخ الاحتياطية. وقد كشف بحث جديد أنّ نماذج تعليم الآلة الخاصّة بجرعات الأدوية الشخصيّة تقوم بتسريب العلامات الوراثيّة للمريض، حيث إن مجموعة صغيرة فقط من إحصائيات الأمراض والمورثات تكون كافية للقراصنة للتعرّف على شخص معيّن، على الرغم من آليات الحجب المتوفرة.

وبطبيعة الحال، المستخدمون الذين لا يشعرون بالرضا عن هذه المخاطر المكتشفة حديثاً يريدون لبياناتهم والتأثير الناتج عن النماذج والإحصائيات أن تكون منسية تماماً.

يعدّ الأمان سبباً آخر أيضاً، والذي يتمثل في أنظمة كشف التطفل غير الطبيعي والمستخدمة في كشف البرمجيات الخبيثة. ومن أجل تحديد الهجمات الخبيثة بشكل صحيح، يتوجب على النظام أن يكون قادرًا على التعرف على الأنشطة الطبيعية للنظام، ولذلك فإنّ أمان هذه الأنظمة يتم بناؤه على نموذج السلوك الطبيعي للنظام المستنتج من بيانات التدريب.

ومن خلال تدنيس بيانات التدريب، فإنّ القراصنة يدنسون النموذج (نموذج التعلم) ويعرضون النظام للخطر. وعندما يتم التعرّف على البيانات المدسوسة، فعلى النظام أن ينسى تلك البيانات وأصلها بالكامل بهدف استعادة الأمان.

إنّ الأنظمة المتعلمة المستخدمة على نطاق واسع، كمحرك بحث غوغل **Google Search**، وذلك بنسبةٍ أكبر، قادرة فقط على نسيان بيانات المستخدم الأساسية، ولا تنسى نتائج البحث عن هذه البيانات، ويتم استدعاؤها عند الحاجة لها. ويسبب ذلك مشكلة للمستخدمين الذين يرغبون في تأكيد حذف أي أثر للبيانات غير المرغوبة بشكل كامل. ويشكّل ذلك أيضاً تحديًا بالنسبة لمزودي الخدمة المعنيين بتنفيذ طلبات إزالة البيانات والحفاظ على ثقة الزبائن.

إنّ مزودي الخدمات سيصبحون بحاجة لأن يكونوا قادرين على حذف البيانات ونتائجها بشكل كامل مع وجود قوانين تحكم خصوصيّة المستخدم مثل "الحق بأن تكون منسيًا" والذي صدر الحكم به في العام 2014 من قِبل محكمة الاتحاد الأوربي العليا.

وفي تشرين الأول/ أكتوبر 2014، حذفت غوغل أكثر من 170,000 رابطًا لتطبيق هذا الحكم، ما يؤكّد على حق المستخدمين بالتحكم بما يظهر عند البحث عن أسمائهم. وفي تموز/ يوليو 2015، صرّحت غوغل بأنّها قد استقبلت أكثر من ربع مليون طلب كهذا.

كسر التبعيّة:

بناءً على العمل المُقدّم في ندوة العام 2015 لجمعيّة مهندسي الكهرباء والالكترونيات **Institute of Electrical and Electronics Engineers** اختصاراً **IEEE**، إن طريقة كلٍ من كاو ويانغ "إلغاء تعليم الآلة **machine unlearning**" مبنية على حقيقة أنّ غالب

الأنظمة القابلة للتعلّم يمكن أن يتم تحويلها إلى صيغة قابلة للتحديث بشكل تدريجي دون الحاجة لإعادة تدريبها بدءاً من الصفر.
ويقدّم منهجهم طبقة من عدد صغير من عمليات الجمع بين خوارزميّة التدريب وبيانات التدريب بهدف إلغاء تبعيتهما لبعضهما. وبالتالي، فإنّ خوارزميات التدريب تعتمد فقط على عمليات الجمع، وليس على بيانات الأشخاص.

باستخدام هذه الطريقة، لن يتطلّب إلغاء تعليم جزء من البيانات مع السلالة الخاصّة بها إعادة بناء النماذج والسمات التي تتنبأ بالعلاقات بين قطع البيانات. ولكن ببساطة، إعادة حساب عدد صغير من عمليات الجمع سيكون كفيلاً بحذف البيانات وسلالاتها بشكل كامل، فضلاً عن كونها أسرع من إعادة تدريب النظام ابتداءً من نقطة الصفر. يعتقد كاو وأنّه ويانغ هم أوّل من أسّسوا صيغة العلاقة بين إلغاء التعليم وعملية الجمع.

وقد اختبر كاو ويانغ منهجهما في إلغاء التعليم على أربعة أنظمة حقيقيّة مختلفة: لينس كت **LensKit**: وهو نظام توصية مفتوح المصدر، زوزل **Zozzle**: كاشف عن البرمجيّات الخبيثة مغلقة المصدر بلغة جافا سكريبت **JavaScript**، أو فلتر البريد المزعج على وسائل التواصل الاجتماعيّة على الإنترنت المفتوح المصدر، و**PJScan**: كاشف ملفات **Portable Document Format** اختصاراً **PDF** الخبيثة وهو مفتوح المصدر.

إنّ نجاح عملية التقييم الأوليّة هذه قد وضع حجر الأساس للأطوار التالية من المشروع، والذي يتضمّن ملاءمة التقنيّة مع نظام آخر، وخلق إمكانيّة إلغاء تعليم الآلة يمكن التأكّد منها عبر اختبار ثابت فيما إذا بالفعل قامت عمليّة إلغاء التعلّم بإصلاح النظام أو قامت بحذف البيانات الغير مرغوبة بالكامل.

في مقدمة ورقتهم البحثيّة، يقول كاو ويانغ بأنّ "إلغاء تعليم الآلة" قد يلعب دورًا أساسيًا في تحسين الأمان والخصوصيّة، وفي مستقبلنا الاقتصادي:
"نتوقع أنّ يتسابق مزوّدو الخدمة على اعتماد الأنظمة القابلة لأن تنسى لأنّها مفيدة لهم وللمستخدمين على حدٍ سواء.

ومع المرونة التي توفرها هذه الأنظمة عند طلب نسيان بيانات معينة، يملك المستخدمون إمكانيّة تحكم أكبر ببياناتهم، كما أنهم سيرغبون بمشاركة بياناتهم بشكل أكبر مع تلك الأنظمة. كما أن المزيد من البيانات تعني إفادة أكبر لمزودي الخدمة، لأنّهم يملكون فرص ربح أكبر ومخاطر قانونيّة أقل".

ويضيفون: "نتصوّر بأنّ الأنظمة القابلة لأن تنسى تلعب دوراً حاسماً في أسواق البيانات الناشئة حيث يتاجر المستخدمون بالبيانات مقابل المال، أو الخدمات، أو بيانات أخرى، وذلك لأنّ آلية النسيان تسمح للمستخدم بأنّ يلغي صفقة بيانات أو تأجير حقوق استخدام بياناتها بدون التخلي عن ملكيتها".

- التاريخ: 2016-05-26
- التصنيف: تكنولوجيا

#Artificial intelligence #algorithm #security

المصادر

- sciencedaily

المساهمون

- ترجمة
  - محمد اسماعيل باشا
- مُراجعة
  - كنان حاضري
- تحرير
  - أنس الهود
  - بنان محمود جوابره
- تصميم
  - Tareq Halaby
- نشر
  - أنس شامي