

¿Cuándo usar el aprendizaje automático?

Bueno, acabamos de hablar sobre varias maneras diferentes en que las redacciones han estado utilizando el aprendizaje automático, y espero que haya sido inspirador.

Pero déjenme intentar darles una guía sobre cuándo deberían considerar usar o no el aprendizaje automático para un proyecto.

La primera pregunta que deben hacerse es: para lo que sea que estén tratando de construir un modelo, ¿pueden simplemente hacer esta tarea manualmente? Por ejemplo, supongamos que están analizando crímenes. Tienen 300 informes de diferentes delitos que deben clasificar. Se necesitará mucho más tiempo para construir un modelo que pueda clasificar estos ejemplos y probablemente necesiten más datos que el que les tomaría a ustedes, o a ustedes y sus pasantes clasificar estos datos manualmente. Si es posible que puedan hacerlo todo de una vez, siempre prefieran realizar su tarea manualmente.

La siguiente pregunta es: ¿tienen ustedes el conjunto de datos que necesitan para hacer lo que quieren? Hay muchas tareas generales en las que pueden usar un modelo preexistente. Hablaremos de esto un poco más adelante en la sección de herramientas. Pero si solo tienen una imagen y quieren saber qué hay en la imagen, ¿es una estatua? ¿Una silla? ¿Un perro? – Ustedes no necesitan un gran conjunto de datos de entrenamiento, probablemente puedan usar un modelo existente para hacer esto. Pero si desean hacer algo muy específico, como identificar aviones espía, necesitarán cientos o miles de ejemplos de vuelo y sus etiquetas. Por lo tanto, ustedes o tienen que tener estos datos, saber dónde encontrarlos o estar dispuestos a invertir en la creación de este tipo de conjunto de datos.

Siguiente pregunta: ¿necesitan que la IA sea explicable? ¿Qué es que la IA sea explicable? Creo que mucha gente entiende que cuando un modelo de aprendizaje automático, especialmente una red neuronal, hace una predicción, no siempre entendemos cómo se llegó a esta conclusión. Por ejemplo, si un delito se clasifica como un delito violento, ¿qué palabras analizó el modelo para llegar a esta conclusión? Ya saben, no siempre está claro incluso para el creador del modelo. A veces no importa si entienden por qué un modelo tomó una decisión. Porque, por ejemplo, en el caso del L.A. Times, cuando estábamos tratando de ver si un crimen se clasificaba erróneamente como violento, los periodistas y editores podían analizar y decir: “OK, este modelo hizo esta predicción, pero déjenme que compruebe si esto es correcto”. Entonces, el modelo les está ayudando de alguna manera, pero el editor siempre está ahí revisando los resultados. Pero deben considerar si al reportar una historia, es importante saber qué hizo el modelo porque es posible que no tengan una explicación satisfactoria de por qué el modelo tomó una decisión en particular.

Y finalmente, y lo más importante, ¿pueden tolerar el error? Prácticamente todos los modelos de aprendizaje automático cometen errores. Es inevitable. Y la forma correcta de utilizar el aprendizaje automático es abordar estos errores, comprender cuándo son realmente serios,

comprender cuándo son aceptables y comprender cómo tratarlos. Entonces, lo que he enfatizado es que ustedes realmente desean utilizar el aprendizaje automático para complementarlo con el trabajo de reportería. Y pienso en eso que el uso del aprendizaje automático es como para encontrar agujas en los pajares o para encontrar pistas en medio de montones y montones de datos, de modo que un modelo les diga “es posible que deseen investigar este avión” o “es posible que deseen investigar si este crimen fue violento o no”. Entonces ustedes trabajan con el modelo, pero saben que cometerá errores. Así que simplemente cuenten con ellos en su proceso de reportería.

Espero que haya sido un buen resumen de cuándo deberían y no deberían usar el aprendizaje automático.