

O algoritmo do BuzzFeed que revelou aviões espiões secretos

Resumo por Adriano Belisário da reportagem [‘We Trained A Computer To Search For Hidden Spy Planes. This Is What It Found’](#)

Para investigar o que se passa sob os céus dos Estados Unidos, o BuzzFeed recorreu aos dados e analisou informações de 20 mil aviões, coletadas do site Flightradar24 durante quatro meses. Os jornalistas criaram um algoritmo de aprendizado de máquina (machine learning) que os ajudou a revelarem diversas operações de aeronaves militares.

O trabalho rendeu algumas reportagens. A investigação revelou atividades de um avião espião secreto norte-americano utilizado no combate ao narcotráfico mexicano e de um avião de vigilância que sobrevoa os céus dos Estados Unidos, entre outras informações.

Inicialmente, foram feitos alguns cálculos com os dados dos vôos de cada uma das aeronaves. Eles incluíam informações como suas velocidades, altitudes, o trajeto de cada vôo ou ainda suas durações. Também foram considerados dados das aeronaves em si, como o fabricante e modelo.

Depois, a reportagem identificou previamente - de forma manual - 100 aeronaves oficiais do FBI e do Department of Homeland Security. E depois selecionou 500 aeronaves aleatoriamente dentre os dados.

A partir desta classificação inicial, o BuzzFeed utilizou um algoritmo de aprendizado de máquina (machine learning) chamado Floresta Aleatória (Random Forest). Basicamente, ele cria um modelo para representar cada um dos grupos anteriores (aeronaves oficiais ou aleatórias). Não por acaso, como os aviões espiões voam em círculos, o algoritmo identificou as taxas de rotação das aeronaves como um fator importante.

Então, os repórteres utilizarem este modelo de classificação para lidar com os outros milhares de avião. O algoritmo não foi infalível. Atividades de paraquedismo foram indevidamente sinalizadas, por circular em uma área relativamente pequena, assim como as aeronaves de vigilância, que foram utilizadas como base para “ensinar” o sistema.

De todo modo, graças à análise, o BuzzFeed conseguiu descobrir rotas de aviões militares e das agências de combate ao tráfico de drogas, além de aeronaves do governo registradas em nome de empresas fictícias e um avião que realiza escutas de rádio nas fronteiras. Todo [código \(em R\)](#) utilizado durante a apuração foi disponibilizado.