

为何千万不能和大数据分析公司合作

作者: linker

原文链接: https://ld246.com/article/1578623737866

来源网站:链滴

许可协议: 署名-相同方式共享 4.0 国际 (CC BY-SA 4.0)

为何干万不能和大数据分析公司合作

最近收到很多大数据公司深度数据合作的案子。大部分价格还不便宜。

但是,这些合作的最大的风险,并不是浪费几十万几百万资金,而是,导致数据精髓的泄露。

等等, 等等, 合同里, 不都是会说, 绝不会把客户数据给第三方使用么?

是的,因为大数据分析公司没必要这么做。

数据就好比是咖啡豆。每家公司都有自己的口味的咖啡豆。大数据公司的会鼓吹,只要贵公司合作, 们就能做出可口的混合咖啡。而且,贵公司的咖啡豆,我们一粒也不会以偷。

听起来完美?

并不是。大数据处理,类似于煮咖啡,煮过的咖啡豆可以保持完整;大数据公司拿走汤汁即可。

在大数据领域,煮咖啡过程,就是用各种模型在数据上训练,一些深度神经网络模型,高达100+层训练完成后其大小高达500+MB。在模型参数的几百兆字节里面,一个原始数据都没有。但是几乎全的数据拓扑都被模型学习和掌握。

大数据公司通过整合各个公司的模型,而不是通过盗用各个公司的数据,一样可以获得整合全部数据效果,甚至效果更好。(因为模型更泛化,体积小,更方便使用)

这种行为可能并没有违法也没有违反数据合作的协议,但是正如我上面打的比方,你们的咖啡豆一粒少,都是被人煮过了,损失还不大么?

原文链接: 为何干万不能和大数据分析公司合作