**隐私保护的机器学习入门参考书目**

# 前言

* 机器学习算法是人工智能（AI）的基础。
* 机器学习分为**三个阶段**：数据收集和预处理、模型训练、模型验证
* 安全机器学习粗略包括**两个大块**：

1）隐私保护：隐私保护的机器学习训练、隐私保护的机器学习测试（oblivious predictions）。

2）安全：各种各样的针对机器学习三个阶段的攻击，例如针对数据收集阶段的数据投毒攻击；针对模型训练阶段的后门攻击；针对模型验证阶段的成员推理攻击。

由于我的方向是隐私保护的机器学习，下面介绍中会侧重于“隐私保护的机器学习训练”这个方面。

# 机器学习入门

* 学校有开一门“人工智能”的课，刘峡壁老师的，了解一下大概的概念挺好的，推荐。
* 线性代数是基础，推不动公式可以回去看看本科的线性代数的书
* 离散数学对写论文的时候写公式很有帮助
* 其他入门书籍推荐：《周志华的西瓜书》、《统计学习方法》。看书顺序是先西瓜书、再统计学习方法

文本

描述已自动生成

# 联邦学习入门

联邦学习是一个机器学习训练框架，现在很火。但它还能火多久未可知。由于我所学有限，在我有限的认知中，没有书籍推荐。

首次接触联邦学习可以去谷歌搜一搜联邦学习的帖子，了解其大概工作流程。其实很简单，几句话就能概括了：在联邦学习中，训练数据在不同的机器；每台机器从服务器下载全局模型，然后利用本地数据训练全局模型；之后返回给服务器本地参数；服务器聚合各机器上的返回的参数，更新全局模型，再把最新的模型反馈到每台机器，完成一轮模型迭代。

建议直接看论文会更清晰，下面列一些联邦学习入门论文。论文的顺序代表推荐观看顺序：

* 综述2019-Advances and Open Problems in Federated Learning（这论文十分推荐，很多联邦学习大佬一起写的，非常厉害）
* 平均聚合函数-Communication-Efficient Learning of Deep Networks（联邦学习始祖）
* CCS2016-Deep Learning with Differential Privacy（联邦学习结合差分隐私，很经典）
* 综述2019-Federated Machine Learning Concept and Applications
* 2020-A Survey on Federated Learning Systems
* Federated Learning in Mobile Edge Networks A Comprehensive Survey

# 隐私保护机器学习

要开始学习隐私保护这个方向，自然需要了解一些密码学知识。下面两本书基本课程，推荐值5颗星，可以日常备一本，随手翻翻：

* Introduction to Modern cryptography(second edition)
* 现代密码学原理（其实就是上面这本书的翻译）

文本, 信件

描述已自动生成

我看过的比较经典的隐私保护机器学习论文，有很多，下面列出一些，不够看的可以再找我要：

* 14NDSS-Machine Learning Classification over Encrypted Data
* CCS17-Practical Secure Aggregation for Privacy-Preserving Machine Learning（联邦学习的隐私保护相关）
* 17S&P- SecureML A System for Scalable Privacy-Preserving Machine Learning
* 17TDSC-Efficient and Private Scoring of Decision Trees, Support Vector Machines and Logistic Regression Models based on Pre-Computation
* 17pattern recognition-Training Support Vector Machines with privacy-protected data