



机器学习中的 Linux 使用简介

赵之旭
人工智能与机器人研究所



目录

- 为什么科学研究中大家大量使用Linux
- Linux的文件系统简介
- Linux的常用命令
- Linux下安装软件
- Linux的环境变量
- Linux下的Python和机器学习框架

为什么科学研究中大量使用Linux

- Linux是开放的、自由的、免费的，大量优秀的程序库出现在Linux上，大量优秀的科学应用软件优先支持Linux
- Linux系统可自由定制
- Linux下编写程序十分方便
- Linux拥有深厚悠久的黑客文化

Linux 文件系统简介

/	根目录。所有的目录都是从这里开始。
---/bin	存放了标准的linux工具，比如像ls、vi、mv。
---/etc	存放了系统配置文件，系统环境变量在这里配置。
---/home	用户目录
---/usr	存放了额外的工具，包管理器管理此处
---/usr/bin	可执行程序
---/usr/lib	库文件
---/usr/share	共享数据
---/usr/local	手动安装的工具，包管理器不管理此处，与/usr结构类似
---/opt	独立的程序，环境隔离

Linux下安装软件

- 包管理器安装
Ubuntu的apt, Redhat的yum, Arch的pacman
一般放在了/usr里对应的/lib, /bin, /include中
- 编译安装
编译前可配置安装目录, 一般应在/opt或/usr/local
- 直接解压
解压到某个目录

Linux中的基本常用命令

- `ls` 列出当前目录的文件，`-a`列出包括隐藏文件，`-l`列出详细信息
- `cd` 改变当前工作目录
- `cp` 复制文件，`-r`递归复制文件夹
- `mv` 移动或重命名文件
- `mkdir` 新建目录
- `rm` 删除文件，`-r`递归删除文件夹，`-f`强制删除
- `man` 查看用户手册（非常有用）
- `su` 切换用户，`sudo`暂时使用root用户执行

Linux中的高级常用命令

- git 流行的版本控制软件，可以参考我写的教程，持续更新中：<http://139.199.163.114/git-tutorial/git-tutorial.html>
- apt 软件包管理工具，用来安装、卸载、升级软件
- grep 使用正则表达式查找文本
- top, htop 进程管理、负载管理，类似于任务管理器
- kill 杀死进程
- nc 简单的“QQ”，在电脑间传输小数据很方便

你需要选择一个编辑器

- gedit, 系统自带的简单小巧编辑器
- emacs 伪装成操作系统的编辑器, my favorite, 强大快速的文本编辑功能, 支持几乎你能想到的各种语言, 加以配置可以支持各种代码跳转、文档查看、编译执行、调试、听歌、看网页。。。
- vi, vim 与emacs齐名, 也十分强大, 系统自带。以上两款学习曲线较陡, 但学会了受益终生
- sublime-text, 一款很出名的闭源编辑器

Linux的环境变量

- PATH

可执行程序的路径，终端命令搜索此路径来执行程序

- PATH的存放文件

`/etc/profile`

系统级环境变量

`/home/<user>/.bashrc`等

用户级环境变量

- 使环境变量生效方法

`source` 命令，例如 `source /etc/profile`
或者重新登陆

Linux的Python和机器学习框架

- Linux的Python是随系统自带的
- 使用python-pip安装python库
`sudo apt install python-pip`
- 流行机器学习框架
 1. tensorflow
`sudo -H pip install tensorflow`
tensorflow-gpu版本需要先安装cudnn，参考官网
 1. mxnet
 2. caffe
 3. torch
 4. keras