预习报告

电气2011-U202012638-龚楷程

1. 预习目标
   1. 对课程的期望

我之前只学过学校开设的C++与Audino课程，除此之外就没有编程的基础了。对于机器学习，当初也没有选上建模课程，所以只是有一些概念的理解。选择这门课是希望能借着这个机会提升自己利用编程工具解决实际问题的能力。先前遇到一些和同学们一起做项目的情况，在分工时我总是会偏向于选择做实物部分，因为软件编程一直是我没怎么接触的，比较薄弱的方面。我有一个学计算机的同学会常常跟我讲一些树形算法，计算机原理、虚拟机之类的东西，但是我对于编程的认识仅局限于在给定框架下写出特定的语句，比如说学c++的时候我知道要写主函数main，然后在下面写代码才能跑，但是不明白为什么。经过这门课，我希望我能：

1. 学习利用python进行编程。一直听说python适用于数据分析，以往遇到这方面的内容总是求助于其他同学，想通过这个机会学习python语言与利用python进行的数据分析。学会这项技能，以后不用每次都去找人家。

2. 对于实际编程的操作有些一些尝试与了解，例如建立虚拟环境进行编程，利用github等辅助编程。

3. 有一个关于编程的初步认识，比如python与c的区别，他们的工作方式等。

4. 实际操作使用机器学习，学习使用机器学习的一般操作，了解机器学习的原理与基本思想，能用机器学习的方式解决工程上遇到的问题。

* 1. 对人工智能的理解

提到人工智能，自然而然就会想到阿尔法狗与无人驾驶。百度上对于人工智能的定义是：人工智能是研究使计算机来模拟人的某些思维过程和智能行为（如学习、推理、思考、规划等）的学科。约翰·麦卡尼在1955年提出：“原则上，学习的每一方面或智力的任何特征都可以被准确描述，由此机器才能进行模拟。人们将探索如何让机器使用语言，形成抽象概念和理念解决现今留给人类的问题，并进行自我完善。”换句话说，AI便是用机器复现人的智能，去解决以往那些由人来处理的问题。比如说无人驾驶便是利用机器展现出来的“智能”去取代人类司机在驾驶中的位置。这样做的意义在于，也许富人仍然可以选择自己的人类司机，但我们更多的普通人有了选择共同“雇佣”一位人工智能司机的可能。总的来说，人工智能就是让机器学着人的样子干着人的事。

那么接下来的问题就是，机器怎么学？我的一个计科同学和我讲过他做的一个井字棋项目。让机器不断对弈，寻找到最优的下法。最后他和我说他的模型应该是遍历了所有可能的下法。然后我问他围棋也是这样吗。他说围棋的可能性太多了。最近我去搜了下AlphaGoZero，他的训练方法也是不断地自我博弈。DeepMind公司的创始人说：“阿尔法狗不光是一个下围棋的专家系统，更是机器学习技术、寻找取胜方法的尝试。”就我目前的理解，人工智能的学习方法，也是机器之余人类的最大优势，便是大量的训练。机器有能力遍历更多的可能，找到人类可能见都没见过的解决方法。这些方法以往受制于人的精力等，难以被发现。即使发现了，也可能被称作“窍门”、“卡bug”，但是人工智能可以通过海量的、低成本的尝试发现这些“窍门”、“bug”，并将其作为解决问题的普遍方法。在这个意义上，也许人工智能能超过人类，做的更好。

也是伴随着我对人工智能的了解，对于一些算法的尝试，我理解了大数据的重要性。人工智能能大量训练的另一面便是其需要海量的数据来支持训练。这个大数据就是人工智能的“训练集”。当厂商拿到了比如说我们消费记录，便可以此训练模型研究用户喜好，推广广告或指导下季度生产，由此创造大量的经济效益。

* 1. 预习计划

1. 学习利用anaconda进行虚拟环境搭建，了解anaconda中各种库的下载方法，尝试利用jypyter进行编程，完成python环境的搭建，了解github的使用，同时了解编程过程中会用到的其他操作及其作用，如设置镜像源等。预期能完成上述工具的基本操作，为后续操作做好准备。

2. 初步了解python、c等语言的特点、了解一些编程的基本要求。

3. 简要学习一下《Python基础教程》这本书，学习一下基本的python语言，先以c++为模板看一看一些常用语法的不同，在此基础上尝试一些在c++上从未进行过的操作，如绘图等。预期能利用python进行基本的编程。

4. 了解一下机器学习的基本概念以及常用方法，计划是看一下吴恩达的机器学习，至少看一下前几集对于机器学习的介绍，预计能对机器学习有一个初步的认识。同时尝试简单的机器学习操作。

5. 学习一下tensorflow等机器学习常用架构。

5. 整体目标是在以往编程基础上，对于课程要用到的知识有一个基本的认识，一定程度上补足自己的不足，做好迎接后续课程的准备，并期望能在课程的操作中不断训练。

6. 预习项目分为两项

第一项是基于我在做的一个大创项目——利用探测雷达探测地底空洞。希望能用学习的知识处理一下我遇到的问题。通过遍历得到一定工作情况下探地雷达的最佳频率，并在此基础上研究最佳频率的分布关系。

第二项是我尝试构建的房价预测模型，我尝试了线性回归与神经网络两种做法，因此在提交的作业里有两段代码。

1. 基础知识预习

2.1工具准备

在电脑上安装了anaconda，并建立好环境。利用pycharm与jupyter进行了初步编程的尝试。注册了 GitHub 账号并且学习尝试了git的操作。

2.2语言学习

学习完了《Python基础教程》的前六章节，学习了python语言的条件、循环等语句；学习了函数定义与使用，类的定义与使用；学习了对于列表字符串的操作，如索引，切片等；学习了大量的函数，对于python都有哪些做法有了基本的印象。

除了书之外，我还利用csdn等学习了一些常用的库如panda等的用法。有了一定python编程的基础，能够基本满足课程的要求。

2.3背景知识

除了一些具体的用法，我还去了解了python的特点。明白了python作为解释性语言与编译性语言如C，C++等的区别，了解了python语言的运行方式与其特点。

2.5库与框架学习

我尝试了python一些常用库的学习，如numpy、panda等，在这些库的帮助下能够较为容易开展数据的分析与处理。我还学习了sklearn与tensorflow等，利用他们尝试了简单的机器学习操作。对于机器学习有了基本上的了解。

2.4机器学习与实践

我学习了吴恩达的机器学习课程，也看了一些机器学习的文章，对于机器学习的一些概念有了初步的了解，比如机器学习的发展现状与应用、监督学习与非监督学习的区别、各种不同机器学习及分类。初步了解了深度学习，对于不同神经网络的算法有了初步认识。比如说课程里面通过一个房价预测的模型讲明了神经网络中神经元的意义，讲述通过模拟人眼提取特征的方法建立卷积神经网络。这些课程与文章构建了我对于深度学习的概念。

我了解了一些简单机器学习的算法与原理，对于机器学习的运行机制有了更深的理解与体会。

我从 Kaggle 上下载了一个房价预测的数据集，并尝试了线性规划与神经网络两种方法对于房价进行了预测。能够进行简单的机器学习操作。

1. 预习项目设计

* 项目一：探地雷达的频率规划
  1. 项目选题

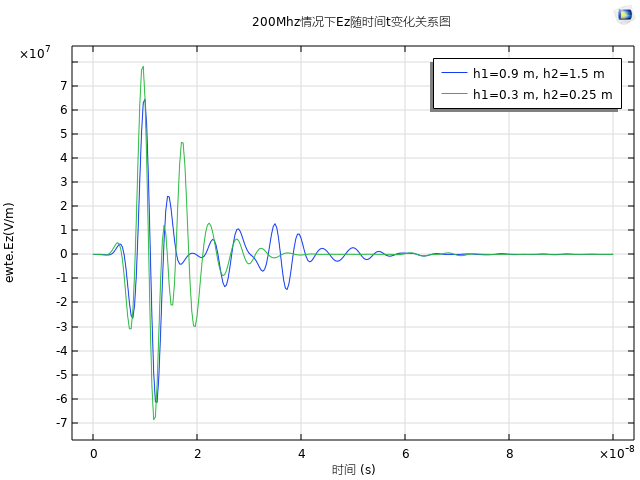
1. 选择这个项目是基于我正在做到一个大创项目——利用探测雷达探测地底空洞
2. 项目功能是基于一些限制条件，用打分的方式为不同的工作环境（地层厚度h\_ground，洞深h\_hole）选择合适频率（bestf）
3. 项目利用python for循环遍历所有工作环境组合（地层厚度h\_ground，洞深h\_hole），在每个组合中遍历100M-1200Mhz频率从而找到该组合下最佳频率，并由此得到一定工作范围内的最佳频率

用循环中存储的bestf绘制最佳频率分布图

用查找表的方式输出某一工作条件组合的最佳频率

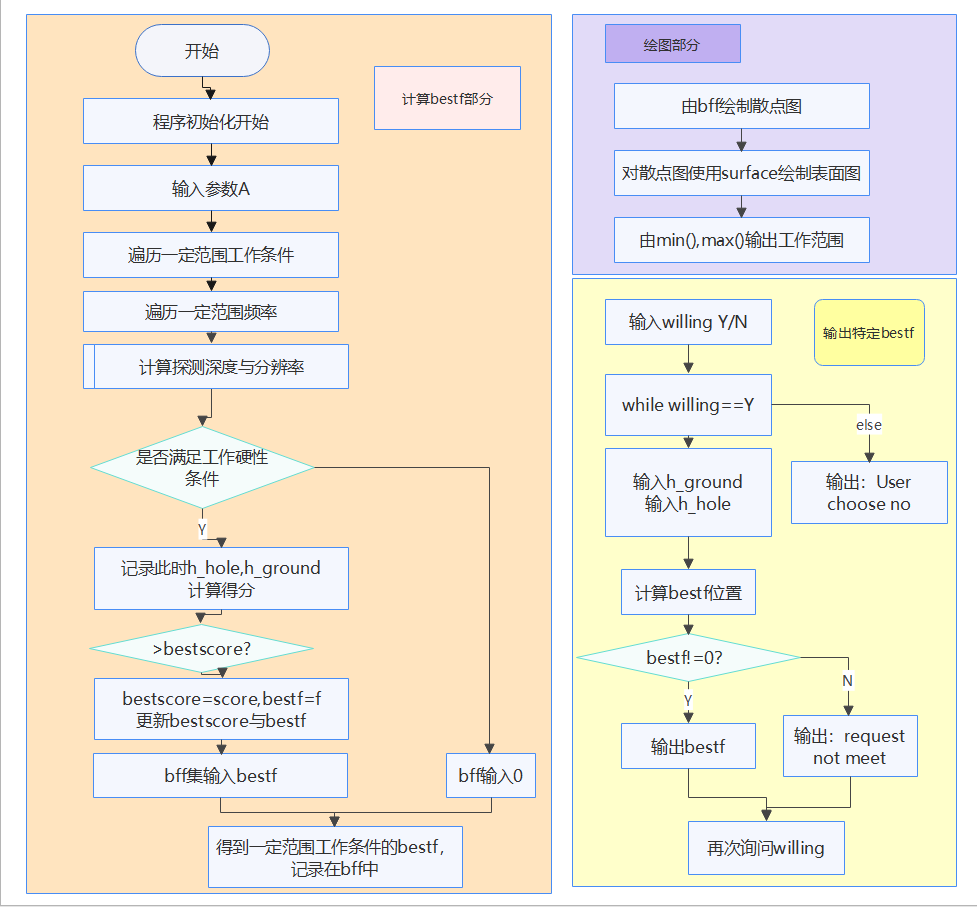
附一张仿真探测波形图：

图中绿色波形便是由于频率选择不合适造成的波形重叠失真，项目的目的便是寻找合适的频率以获得如蓝色波形的清晰结果。



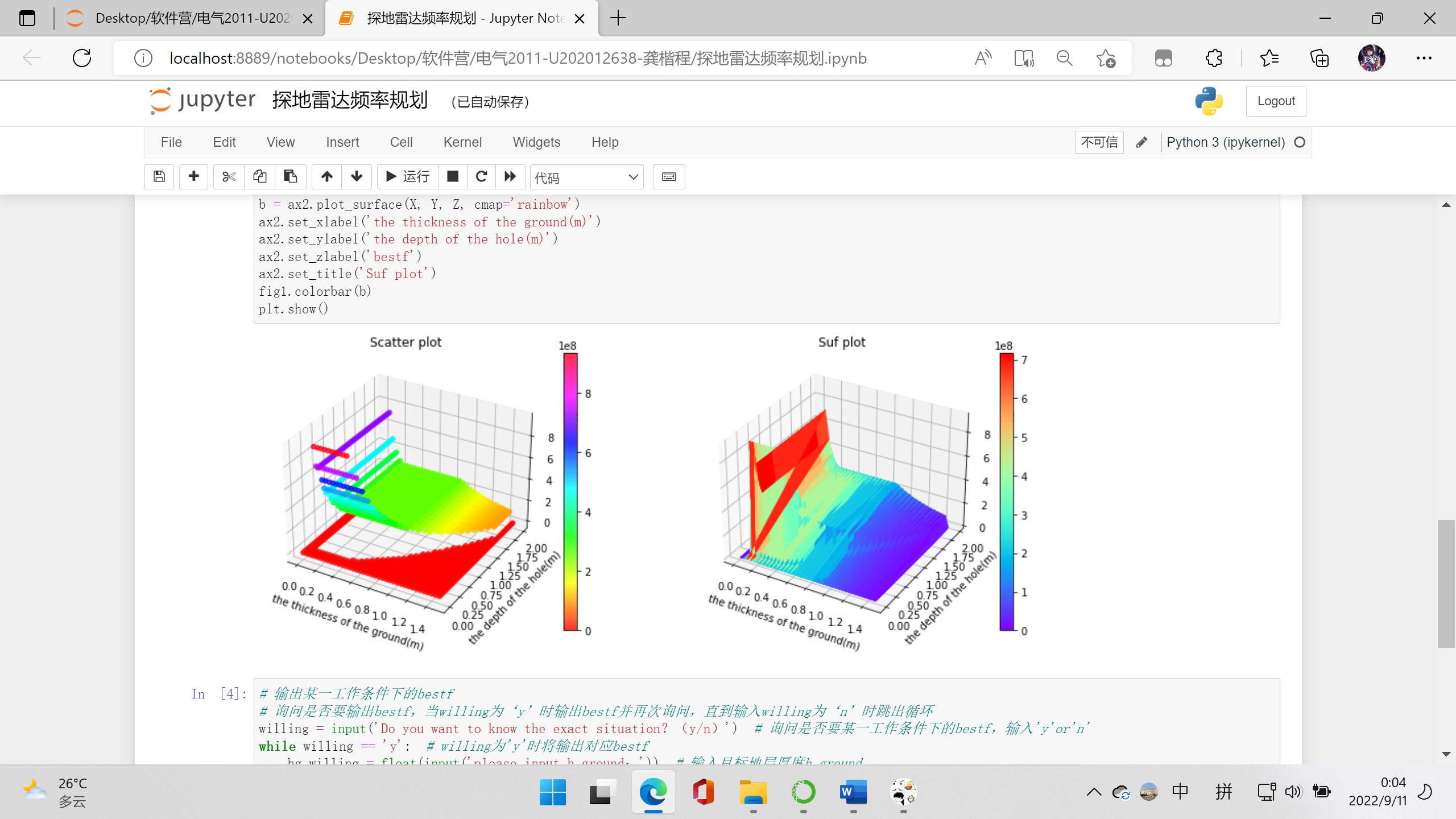
图一 演示波形图

* 1. 程序设计

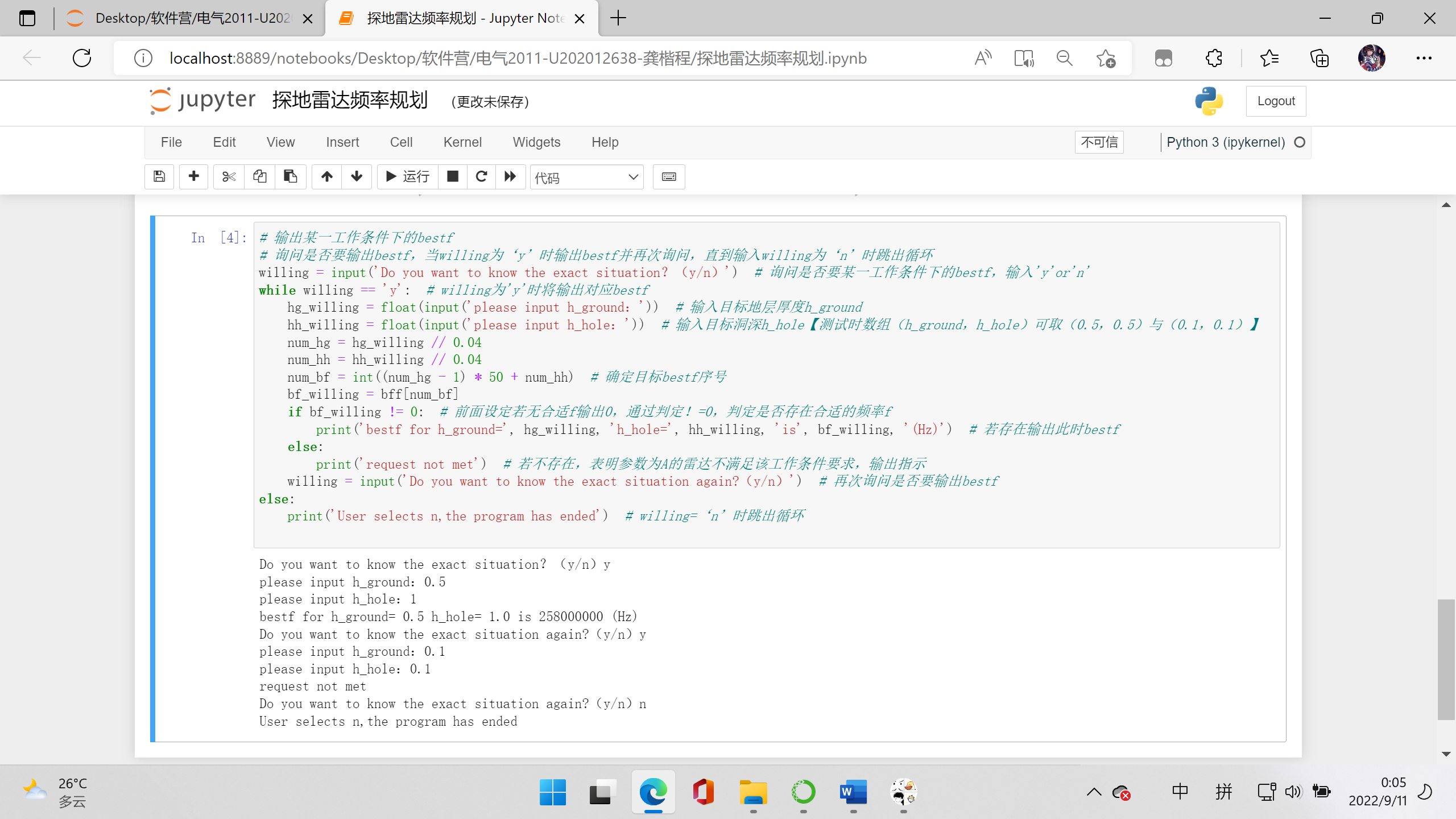


图二 探地雷达的频率规划流程图

* 1. 运行结果
     + 1. 得到了一定范围内最佳频率的分布图，绘制了散点图并利用surf得到了表面图（效果不是很好，可能划分过粗）
       2. 能够查找任意一工作条件的最佳频率，能够为大创项目的进行提供指导



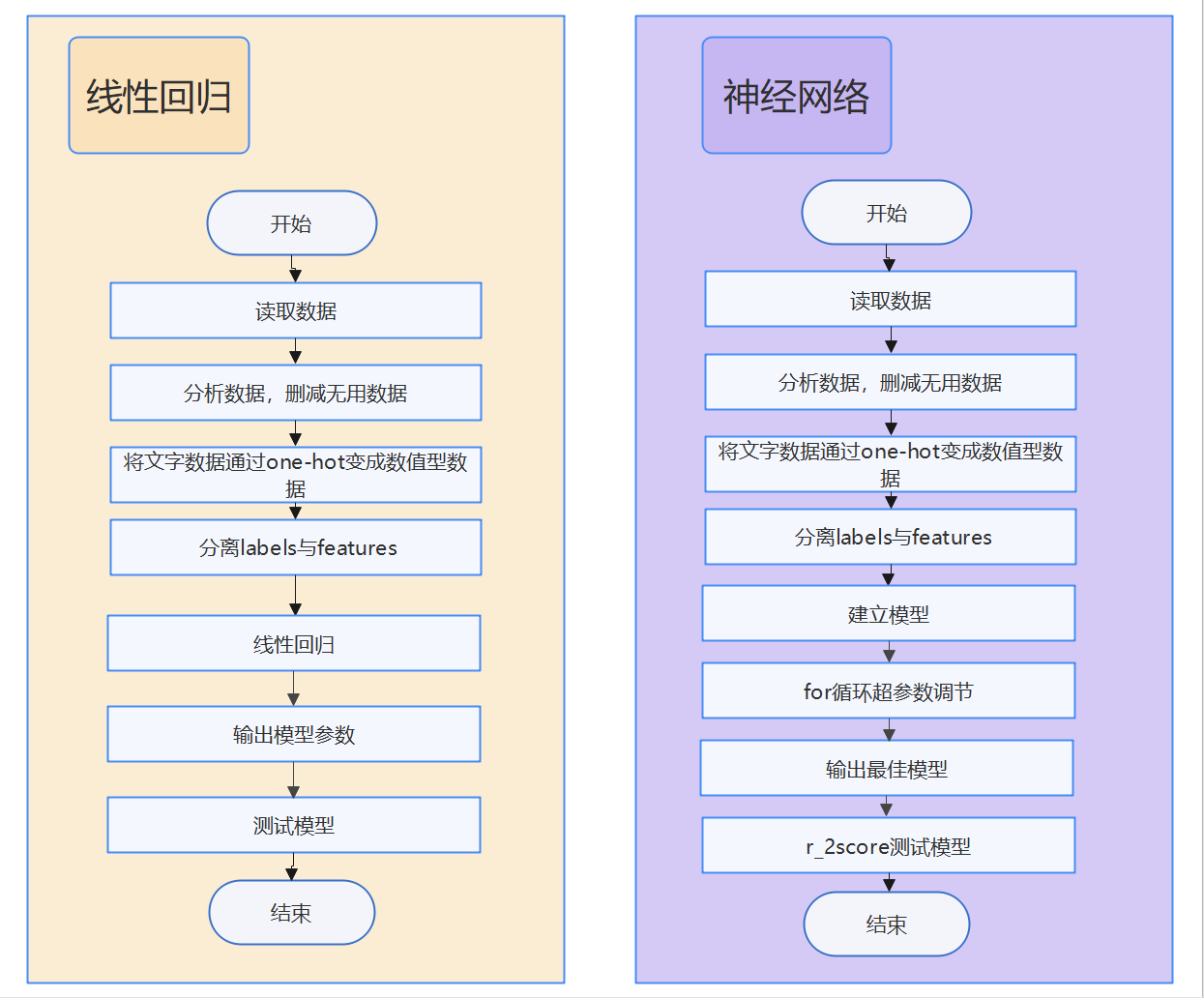
图三 频率分布结果图



图四 .输出某一工作条件最佳频率运行结果

* 项目二：房价预测

1. 项目选题
   1. 这个项目是我在 Kaggle 上找的一个房价预测数据集
   2. 根据给出房子信息预测房价
   3. 尝试分别使用线性回归与神经网络进行预测
2. 程序设计



图五 房价预测流程图

1. 运行结果

分别使用线性回归与神经网络进行预测，用r2\_score对预测结果打分

两个模型都能达到50%以上正确率。对于数据处理等方面应该还有优化的空间，希望通过后面的学习，进一步提升能力，优化模型实现更为准确的预测。

图形用户界面, 应用程序

描述已自动生成

图六 线性回归预测结果

图形用户界面, 文本, 应用程序

描述已自动生成

图七 神经网络预测结果

1. 预习总结

1. 在预习过程中，我初步学习了python的相关语法和操作，对于使用python编程有了基本的概念，能够用python解决实际生活中遇到的一些简单的问题，如简单的逻辑操作，数据的分析、处理以及运算、图形的绘制等，可以为未来项目的运行提供帮助。

2. 我学习了神经网络的相关概念与算法原理，对于神经网络的运行机制有了基本的认识。进行了简单的预测操作，分别使用线性回归与神经网络实现了房价预测。虽然正确率还不是很高，但是通过这些尝试我明白了机器学习的基本流程，了解了数据处理的相关算法如相关系数矩阵，了解了一些回归算法与其相关操作，学习了神经网络的搭建，尝试了调参等操作并利用for循环的方式辅助调参。经过这些预习，我对于机器学习有了更深的理解、对于机器学习功能的强大有了更深的感触，希望通过更深的学习未来能够利用机器学习解决实际问题。目前在初步学习后我想对于我做的第一个预习项目，选择合适的频率是针对于现有常规信号处理方法，由于波形之间有联系，采用机器学习的方式经过多次训练是否能够分辨出一些已经失真的重叠波形。我希望在学习机器学习后未来能够尝试用机器学习来解决这些问题。

经过预习，我学习了python的基本操作，以后在做集体项目分工时可以尝试承担与编程相关的操作，未来想继续学习利用python帮助数据分析处理，能够完成一些简单的项目。学习了基本的机器学习，有了初步的理解，能偶进行简单的机器学习构建，希望能够通过进一步学习将机器学习应用于更多方面，为一些项目提供帮助。