### 自我介绍。

你好，我是王玉联，来自中南大学。现在跟从导师研究图神经网络，我主要是做图神经网络在不同场景下的应用，包括ACM源代码挖掘和恶意软件的静态检测。

ACM源代码挖掘项目，我们提出一种新的特征试图从语法、语义和宏观的函数调用这三方面理解程序，使用图模型分类的精度在97%以上。在类别非常相似的分类任务中，相对于TBCNN和GGNN作，精度分别提高了9%和4%。

关于恶意软件的静态检测这个项目，我们主要是考虑到当前的恶意软件检测的方法虽然精度都很高，但是不会很明确的告诉用户为什么会被检测为恶意软件。因此，我们在这个项目中使用基于注意力的深度神经网络，挖掘对模型分类有较大影响的敏感API序列，并通过这些序列去反向的分析可能涉及到的恶意行为，具有很好的可解释性，而且分类精度达到97%以上，比集成算法高4%左右。

我面试的岗位是测试开发岗。因为我本身有过程序开发的实习经验，具有一定的编程能力，而且在工作上也是一个比较认真和细心的人，所以我认为自己比较适合这个岗位，也希望自己能够在这个方向上发挥自己的潜力。

### ACM源代码挖掘项目

中，我们提出一种新的特征从语法和语义两方面理解程序，并使用GGANN模型进行程序分类，精度在97%以上。

**分工：**融合三个图的特征、使用和改进模型是共同讨论的结果，然后另一个同学是负责进行在AST的基础上进行加边，我是负责改进模型的部分。

**为什么融合：**考虑到各个图的特点，抽象语法树可以刻画程序的语法结构，而数据流图表示数据在程序中传递和加工的过程，函数调用图表示的是更为宏观的函数之间的调用关系，我们希望能够将这些信息融合到同一个图里面，这样这个图就会拥有更加丰富的信息。

**如何融合的：**主要方式是在抽象语法树的基础上加入不同类型的边。我们认为抽象语法树上的边是一种类型的边，然后在此基础上加入函数调用边。数据流图上的边比较复杂，有不同的语义信息，所以我们将数据流图上的边大体分成几个部分，有计算边、返回边、参数边等等。举几个例子吧，比如说a=a+1;，等号左边的a和右边的a就存在一种计算的关系，一个函数中return语句返回的变量和这个函数之间就会存在返回边，参数边就是在一个函数中的实参和形参之间存在着的边。最终形成的图中共有7中类型的边。

**使用和改进的模型：**我们使用了GGNN模型，利用类似RNN的原理实现信息在图上的传递。RNN比较经典的应用是填槽技术，我理解的是相当于单词填空。就是给定一句话，比如说我爱我的祖国。然后把祖国这个单词藏起来，RNN就可以通过’我爱我的’这一单词序列取预测下一单词的概率，比如是父母的概率是0.11，是祖国的概率是0.15。GGNN模型就是相当于把图的每一个状态看成是一个单词，然后预测图的下一个状态，直到图的状态收敛的时候，模型也就训练好了。我们是在GGNN模型的基础上加入了双重注意力机制，第一层注意力是节点与节点之间的，第二层是节点和图之间的。

**最终结果：**最终改进的模型与原有模型在使用同样特征的情况下，精度提高1%左右。当时我们做的时候，TBCNN是一款比较好的程序分类软件，它是一款基于抽象语法树的卷积神经网络。使用我们提取的特征和改进的模型与TBCNN比较，精度提高接近4%。同时，我们还加大难度，从数据集中选择出非常相似的题目，比如说两数求和这样比较简单的编程题目，一个是对整数求和，一个是对浮点数求和。在这样类别非常相似的分类任务中，我们的工作想对于TBCNN提高9%，相对GGNN提高4%。

**代码量：**模型比较大，加上数据处理代码应该在1000行左右，因为讨论时候的分工是由另一个同学进行特征的提取，我主要负责模型的改进，所以在这个项目中我修改和添加的代码应该在50-100行之间。

### 恶意软件检测项目

中，我是从从函数调用图的角度去考虑，因为函数调用图只能够反映软件的结构特性，如何为函数调用图添加更多的信息成为我们考虑的一个主要问题，因此我们通过挖掘图中节点对应函数的敏感API序列，获取更为细粒度的内容信息，然后通过基于注意力机制的神经网络实现恶意软件的识别和检测相对于集成算法提高4.5%左右，且获取的结果具有更高的可解释性。

**分工：**几乎是我独立完成的，从特征提取到模型训练，除了使用susi提取敏感API筛选列表。

**使用了什么样的模型：**使用的是基于注意力的神经网络模型。在提取的图特征的基础上，使用随机游走的方式获得一些更加长的敏感API序列，并对路径进行编码。随机初始化一个向量作为图的向量表示，计算每一个路径与图向量的相似度，通过模型的前馈和反馈不断调整图向量的值，直到模型收敛。

**可解释性是什么意思：**在模型训练的时候，我们通过计算图向量与敏感API序列的向量的相似度更新图向量，最终学习到图向量。在这一过程中，每条路径的相似度我们对这条路径的注意力，注意力越大，说明这条路径对最终分类的影响越大，也说明了这条路径就越重要。我们可以通过这些重要的路径，分析恶意软件可能产生的恶意行为，比如说发送收费短息、侵犯个人隐私等等。

**代码量：**整个项目在500行左右，有部分比如说随机游走这种，是有代码的，会直接拿来改改用。

### 印象较为深刻的bug:

印象最深刻的可能就是有一个ticket，需要我新添加三个接口，这三个接口之间的增删改查是有相互影响的。要求我三天内转给测试，当时看到ticket的需求说明后就马上开始工作了，但是当我花了将近两天的时间搞完然后请前辈过来review代码的时候，前辈发现功能好像有点过去繁杂了，前辈是一个比较细心和负责的人，就去看了下ticket的需求说明，发现我俩的理解不太一样，所以就找来产品经理确认，发现确实是我想的太过复杂了，实现了过多的功能。通过这件事情我觉得做事之前的考虑真的要尽量细致、多提问、多确认，这样才能尽可能的不去做冗余的事情。

为自己写的代码写过一段时间的junit测试

### 你对测试开发的了解（为啥要选择测试开发）：

有和从事这一岗位的同学了解过，测试开发需要完成开发那边提交的测试任务，还需要开发一些测试工具和平台，有时还会做一些同业务相关的开发。我对这个岗位是比较又兴趣的，而且结合我自己的一个情况，觉得自己和这个岗位匹配度挺高的。因为我个人的理解是，测试开发这个岗位一方面需要逻辑和发散思维，一方面相对于测试而言对编程能力要求要高一些。而我，+优势

### 那你的优势：

我曾参与过ACM集训和竞赛，研究生阶段在搞算法，平时也比较喜欢玩一些剧本杀和狼人杀这样的推理游戏，逻辑思维和发散思维比较好；然后本硕都是计算机专业的，也有过半年多的实习经验，编程能力比较强。

做事很有计划性，以及讲信用，答应别人的事和计划做的事情我都会记下来，并放在显眼的地方，我曾把项目计划作为手机壁纸，时刻提醒自己。

### 你的缺点

**你，（作为一个来应聘本职位的candidate），（你的资历和经验对于这份工作来说），最大的（我需要知道的、会影响工作）缺点是什么 （以及你是否有计划改正它）？**

虽然我有计算机专业背景，也学过一些测试方面的理论知识，但是由于没有在一个实际项目中没有担任过测试这样一个角色，相对来说会比较缺乏测试的项目经验。我准备利用这一年的空余时间做一两个真实的测试小项目，熟悉一些测试工具、丰富实际的测试经验。

我发现自己容易在做一个任务的时候，担心另一个任务的完成进度，这不是一个好现象。后来我想到一个很好的方法，就是在开始一天的任务前，先把任务列出来，看看能完成哪些，并给每个任务分配的一大概的时间，等到晚上就进行清算，这提高了我的工作效率，而且让我的成就感提高了很多。

### 性格

别人的评价：朋友评价我是一个比较有毅力、有上进心的人，还有就是比较值得信赖，比较守信。

自己的评价：自己是一个比较有计划性的人，比较有责任心和守信。

### 如果一两年后，你发现自己对这个行业不感兴趣，你会怎么办?

我会认真的考虑是否换行业这一决策的机会成本，毕竟自己积累的专业知识、在这个行业认识的人等等，绝大多数可能都会被浪费。对现在的我来说，选择测试开发这个岗位是经过认真分析所选择的，我认为自己会在这一岗位上得到发展和成长，我有这样的信心。

### 强势的产品经理（业务在两天后上线，当前任务已经和紧张，产品经理要增加新的需求）：

首先我觉得每个人都有自己的特点，这个产品经理虽然很强势，但我觉得既然他能够成为产品经理，必然有他的专业素养。我会先向产品经理确认该需求的紧急性，如果该需求是紧急的、不上线会影响整个公司的产品竞争力的话，我觉得任务再紧张，也要加班加点完成。如果该需求并非是紧急的，我会向说明当前开发进度的紧张状况，遇到的问题，说明时间上来不及，并说明匆忙的开发该功能可能会得不到有效的测试以及得不到有效的测试所带来的潜在风险，尝试与产品经理协商能够延迟上线。我相信沟通到这一步，绝大多数的情况都能协商一致了，如果仍然不能达到一致意见的话，作为最后的选择，只能去找teamleader做出决策了。

### Fix某个问题并使你得到成长的事例：

在做恶意软件的识别和检测时，我在函数调用图中挖掘函数节点的API序列，但是挖掘出来后发现每个函数节点的API序列太长了，有太多重复的内容，直接使用的话，肯定不能得到，于是就尝试去除冗余API，比如去除高频的API、去除tf-idf较高的API。但是效果都不是很好，于是去找相关文献去看、并和做这个方向的同学讨论，发现了一篇比较适合解决该问题的文献，使用文献中的方法获取了敏感API函数库，并使用这个敏感API函数库筛选API，得到了很好的结果。我觉得做一个工作不能一拍脑门就去完成了，要多汇总各方面的信息，有时站在前人的肩膀上，会让我们能够更快更好的找到解决问题的方法。

应该就是acm竞赛的经历了，在参加acm集训时，各个专业的学生报名的人很多，但是随着集训的强度和难度不断的增长，很多人都退出了，尤其是女生，就只剩下2个人，后来在校赛时我取得名次，虽然只获得三等奖，但它给我带来的成就感要比获得国家奖学金、保研这些事情还要多。

### 如何看待加班？

其实我个人的想法是尽量的不要依赖加班来完成工作，而是想办法提高自己的工作效率。但是大概了解到互联网公司的加班情况较为普遍， 我既然投互联网公司，就还是做好了加班的准备，毕竟大环境如此。我认为如果公司加班制度完善、能够让我得到技术的提升的话，我觉得作为年轻人拼一拼也还不错。

加班不是做样子，不是为了刷部门加班业绩，不搞攀比，有加班费，加班时间在合理范围内。

### 找工作更关注哪些方面

公司人才培养的方式，晋升机制，公司在做的项目内容，然后是城市、薪资待遇，然后由于互联网公司加班情况比较普遍，还会关注加班制度的问题。

### 职业规划

我希望能够通过工作提升工作经验，沉下心来搞技术，希望自己的专业水平能够得到提升，自己的工作能为公司创造价值。

### 你是一个好学生吗?

好学生，成绩，其他方面包括：参加竞赛、学生干部、义务支教活动、社团活动，全面发展。

### 学生生涯最大的提高是什么

我觉得学生生涯一个方面肯定是给予了我大量的知识以及学习的方法，但是我觉得自己最大的提高还是一次次的测验让我能够在压力下继续思考，以及异地求学的经历让我更具有责任感和更加独立，这种独立不只是生活上，还有心理上的。而且在这个过程中我接触到了来自各个地方的伙伴，他们有不同的风俗、属于不同的人群，对我的影响是让我越来越能够换位思考，理解别人。

### 关于经济学：

有听过一些经济学方面的读书音频分享，有一课经济学、小岛经济学等等，看过一部分薛兆丰的经济学通识，收获还是蛮大的，也挺感兴趣的。也乐于能够多去了解和学习一些金融方面的知识。