

# 交底书评估报告

专利申请名称：能感知不确定性的轻量化网络资源效用预估模型

评估关键词：学生 教师 模型 知识蒸馏

查询范围：中国发明专利公开文献、中国发明专利公告文献、中国实用新型专利文献、PCT 公开文献、专利信息服务平台(CNIPR)、中国专利检索系统文摘数据库(CPRSABS)、中文全文库(CNTEXT)、德温特世界专利索引数据库(DWPI)、CNKI 系列数据库

相关国内外专利检索结果(专利申请号、专利名称)：CN202210837954.2 多层次自适应知识蒸馏的轻量化高分遥感场景分类方法;CN202111466873.8 一种基于轻量化预训练语言模型的文本分类方法;CN202211046284.9 一种基于知识蒸馏的遥感变化检测方法;CN202210488306.0 一种基于多模型对抗蒸馏的鲁棒性图像分类方法;CN202110824459.3 基于局部-整体联合知识蒸馏算法的步态识别模型压缩系统及方法;CN202111109745.8 基于多任务学习与知识蒸馏的车辆再辨识方法及系统;CN202011108692.3 图像超分辨重构模型的生成方法、重构方法及电子设备

国内外文献查询结果(文献名称、来源、发表人)：无

该交底书涉及一种能感知不确定性的轻量化网络资源效用预估模型，经评估，提出如下意见：

技术方案具备专利法第 22 条第 2 款规定的新颖性。

本申请技术方案，请求保护一种能感知不确定性的轻量化网络资源效用预估模型，利用知识蒸馏，条件自编码器和深度学习不确定性预估的方法，对轻量化网络资源效用模型利用全部用户历史数据和网络资源对象特征的条件，进行高效，准确的网络资源效用预估，并且能获得预估的不确定性，帮助资源分配平台把控分配质量。

根据国家知识产权局公布的中国专利文献库的初步检索，与技术交底书较为近似的现有技术为：CN202210837534.4 基于知识蒸馏和因果推理的推荐方法及系统

该文献公开了：一种基于知识蒸馏和因果推理的推荐方法及系统。本发明中，首先把训练数据集中所有用户按照敏感属性的高低等分成若干个用户组。再利用所有用户的行为数据训练一个基础推荐模型，继而按照用户分组，利用每一组用户数据，对基础推荐模型进行微调，为每一组用户训练一个教师模型；最后利用所有用户的数据，借助因果推断中的前门调整方法，通过每个用户分组的教师模型获取多个中间表征作为中介，继而利用 Batch 内采样机制和注意力机制，进行多模型多样本信息聚合，并将聚合后的信息蒸馏到学生模型。本发明将因果知识蒸馏技术应用于项目推荐中，相比于普通推荐算法，引入因果建模可以有效提升用户的推荐服务公平性，缓解马太效应。

由上述内容可知，本次申请的技术方案与上述现有技术文献的区别在于：该技术使用条件自编码器的模型，将轻量模型在线推断时无法使用的信息编码到模型的隐空间中，并且在训练的过程中利用知识蒸馏的方法，利用全部信息，以及预训练的教师模型，提升轻量化模型的预估准确度。并且在条件自编码器中引入了深度学习不确定性，将隐空间中编码的缺失信息帮助轻量化轻量化的学生模型与较为准确的教师模型之间的差异，得到轻量化模型预估的不确定性。

相比之下本申请解决的技术问题是：现有的网络资源效用的预估算法主要聚焦于如何更好地利用用户的长历史信息，在利用更复杂且准确的模型预估网络资源对用户的效用的同时，也需要耗费更多的算力和更长的计算时间。然而网络资源的分配往往需要实时产生，因为需要将其实时的推送给用户，过长的计算时间会影响用户的上网体验。轻量化的网络资源效用预估模型为了减少运算量，无法较好利用较多的信息作为输入，即无法使用用户的长历史记录和全部网络资源对象的特征；另外，这样的模型为了更快地计算，模型结构也更为简单，参数量更少。这两者同时会使得预估的准确度下降，和复杂模型的精确度差距较大。

经进一步检索发现，相关领域中公开的参考文献如下，也记载了与上述技术手段相近似、技术效果接近的装置/方法，包括：CN202110088772.5 一种深度神经网络模型可视化解释方法及系统。

其中记载了：一种深度神经网络模型可视化解释方法，包括：给定待解释的输入图片，利

# 交底书评估报告

用变分自编码器生成合成图片；通过深度神经网络模型对合成图片打标签，依据合成图片的标签是否变化将合成图片分为两类，得到合成标签，从而得到合成数据集；在合成数据集上训练可解释模型，将深度神经网络模型的知识蒸馏到可解释模型中；根据训练好的可解释模型，建立输入图片和深度神经网络模型的预测结果的对应关系，进而得到一张显著图，从而实现对深度神经网络决策结果的解释。本发明能够提高深度神经网络模型的可解释性，使模型透明化，使用户更好地理解深度神经网络的决策过程，从而促进模型的实际部署，推动人工智能领域研究的发展。

但该技术方案与上述文献记载的内容相比，依旧存在一定的区别，两者在解决的技术问题和所采用的技术手段，得到的技术效果上相比并不完全相同。因此上述文献并未明确公开解决区别技术特征的技术启示。同时，结合事务所多年专利代理经验，上述申请的创造性有待在修改过程中进一步补充，才能在专利审查过程中进一步突出与现有技术相比的显著区别：建议申请人详细公开所采用的各个技术手段的细节、技术改进点的可选参数设置以及详细实验数据，即专利法要求的意料不到技术效果。事务所在修改过程中根据补充修改内容及实验数据进一步确定是否明确符合专利法第 22 条第 3 款的规定。仅凭目前技术内容，预期专利局在实质审查过程中能够检索处更为接近的现有技术以及相关技术启示。

专利法第 22 条第 3 款规定的创造性是指：从最接近的现有技术和发明实际解决的技术问题出发，判断要求保护的发明对本领域的技术人员来说是否显而易见。当现有技术中给出将上述区别特征应用到该最接近的现有技术以解决其存在的技术问题(即发明实际解决的技术问题)的启示，这种启示会使本领域的技术人员在面对所述技术问题，有动机改进该最接近的现有技术并获得要求保护的发明时，则发明是显而易见的，不具有突出的实质性特点。

综上，本申请技术方案与上述文献相比具有一定新颖性。故评估该申请：

- 1、 所涉技术方案具有一定工业实用性。
- 2、 所涉技术方案预期经详细补充所有技术细节以及实验数据和技术效果，并经修改后预期可以相比上述现有技术文献体现出新颖性。
- 3、 所涉技术方案相比上述现有技术预期经详细补充所有技术细节以及实验数据和技术效果，并经修改后预期可以相比上述现有技术文献体现出一定创造性。

申请建议：

交底书中记载的具体实施方式技术细节以及相关核心操作参数、关键算法以及实施后得到的实验数据和/或模拟实验数据有待进一步详细补充。从而进一步突出该技术与现有技术的区别。

以上初步评价仅基于技术交底书、申请人课题组提供的信息以及事务所初步查询得到的结果，并不代表专利申请经沟通修改后的文本的专利性。国家知识产权局在申请提交后的实质审查中可能由于情势变更、资源、条件的不同而得出与上述评估结论不同的检索报告。

2022 年 12 月

