## 《计算机新技术讲座》报告写作指导书

叶宏庭 （学号：71118415）

**一、报告内容**

|  |
| --- |
| 1.报告信息：讲师贺刘尧，目前是IBM数据科学家。北京工业大学统计学硕士。毕业后任职于IBM，从事数据科学家工作。擅长运用经典数据分析、机器学习、深度学习、统计及数学模型等专业技能，帮助客户洞察数据背后的特征，进而提供解决方案。金融行业NLP、知识图谱专家。本次讲座采用腾讯会议，线上开展。  2.报告内容：随着AI及大数据的发展，人们的日常生活中或多或少都会遇到AI类的产品。其背后的机器学习和深度学习技术为何？在行业中有什么应用？如何运用这些技术开发实践？本次讲座主要分为以下几个部分， 1.深度学习基本理论与发展现状2.主流的深度学习开源开发框架3.深度学习开发环境搭建利器IBM Spectrum Conductor Deep Learning Impact介绍4.深度学习经典模型和应用开发案例介绍的逻辑和洞察应用到商业模式中去，才能真正地体现数据金矿的价值。 |

**二、收获和扩展学习**

|  |
| --- |
| 1.报告收获：了解了主要一些深度学习网络，感知机和前馈网络，Hopfield网络，卷积神经网络CNN，循环神经网络RNN，对抗神经网络GAN。深度学习是此次AI领域耀眼的明珠，吴恩达曾说过，理论上说现在所有的学习模型都能用深度学习模型替代。在开始我们的深度学习之旅前，读者朋友可以将这些框架和概念建立起来，ARGO之后会在深度学习领域对这些模型做更加深入的介绍。RNN ,CNN和GAN这些网络的设计体现了人类的灵感和智慧，而这些可能是机器学习现在还不可及的。  2.扩展学习：在深度学习火热发展的时代，越来越多的流行框架开始被使用，包括Tensorflow、PyTorch等，通过这次的讲座，后续可以继续针对某一种框架进行深入的学习，尝试掌握一些简单的深度学习开发。 目前深度学习主要有以下领域的应用，1.计算机视觉：（1）目标检测，目标检测（Object Detection）是当前计算机视觉和机器学习领域的研究热点之一，核心任务是筛选出给定图像中所有感兴趣的目标，确定其位置和大小。其中难点便是遮挡，光照，姿态等造成的像素级误差，这是目标检测所要挑战和避免的问题。现如今深度学习中一般通过搭建DNN提取目标特征，利用ROI映射和IoU确定阈值以及区域建议网络RPN统一坐标回归损失和二分类损失来联合训练。（2）语义分割，语义分割（Semantic Segmentation）旨在将图像中的物体作为可解释的语义类别，该类别将是DNN学习的特征聚类得到。和目标检测一样，在深度学习中需要IoU作为评价指标评估设计的语义分割网络。（3）超分辨率重建，超分辨率重建（Super Resolution Construction）的主要任务是通过软件和硬件的方法，从观测到的低分辨率图像重建出高分辨率图像，这样的技术在医疗影像和视频编码通信中十分重要。在深度学习中一般采用残差形式网络学习双二次或双三次下采样带来的精度损失，以提升大图细节；对于视频超分一般采用光流或者运动补偿来解决帧图像的重建任务。（4）行人重识别，行人重识别（Person Re-identification）也称行人再识别，是利用计算机视觉技术判断图像或者视频序列中是否存在特定行人的技术。其广泛被认为是一个图像检索的子问题。核心任务是给定一个监控行人图像，检索跨设备下的该行人图像。2.语音识别，语音识别（Speech Recognization）是一门交叉学科，近十几年进步显著。除了需要数字信号处理，模式识别，概率论等理论知识，深度学习的发展也使其有了很大幅度的效果提升。深度学习中将声音转化为比特的目的类似于在计算机视觉中处理图像数据一样，转换为特征向量，语音识别的难点有很多，例如克服发音音节相似度高进行精准识别，实时语音转写等，这就需要很多不同人样本的声音作为数据集来让深度网络具有更强的泛化性，以及需要设计的网络本身的复杂程度是否得当等条件。3.自然语言处理，自然语言处理（NLP）是计算机科学和人工智能领域的方向之一，研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。深度学习由于其非线性的复杂结构，将低维稠密且连续的向量表示为不同粒度的语言单元，例如词、短语、句子和文章，让计算机可以理解通过网络模型参与编织的语言，进而使得人类和计算机进行沟通。 |