老师好，今天我们组开题答辩的项目是风电功率可视化软件开发。

我的答辩将从一下几个部分进行，分别是：开发背景，开发目的与内容，开发路线，项目分工以及后续设想

由于我们组四人中有三人都是同一大创团队的成员，我们便决定从我们的大创项目“基于人工智能的风电功率预测”入手,进行对应的软件开发，接下来我也会简要介绍我们的项目背景

风电作为新能源发电的一种占比较高，且具有随机性波动性的特点，故实现对风电功率的精确预测具有很重要的意义。而国外已有一些成熟的风电功率预测系统及软件，也使我们的开发变得很有必要。

我们的开发内容主要集中在可视化和预测，而可视化也是我们在工程训练营期间开发的重点。

我们的大创项目基于人工智能方法构建神经网络，选取合适的输入特征，建立风电厂发电功率预测模型，实现对风电场发电功率的准确预测。

由此，我们也希望能够开发一款风电功率可视化的软件，帮助用户更好理解风力发电特性，并帮助专业人士直观了解风电功率状况，合理规划风电场选址。

右侧这张图片展示的是风电功率曲线，在以风速为主要影响因素的基础上加入其他输入特征，可以实现更精确的预测。

我们的开发路线主要集中于建立精准模型，设计美观界面和实现多种功能三个部分。

具体的开发路线大概分六个部分进行，首先需要对数据进行收集和处理，以获得训练模型的基准有效数据。之后要建立模型并加以测试和验证，在此基础上进行后续的软件开发，建立优良界面，并与多种功能相连接，满足用户需求。

在数据收集部分，我们找到了某风电场SCADA系统的实时监测数据，以30秒为颗粒度得到了时间跨度为一周共计104万个数据点，其中包含风速，风向角，气温，桨距角等多个影响风电功率的因素。

之后，我们对收集到的数据进行筛选和处理，基于风电功率曲线的规则和密度聚类算法进行清洗，从104万个数据点中提取得到约60万个有效数据点，作为之后训练模型的输入。

我们在建立神经网络模型时选用了五种较成熟且效果较好的模型，选择不同模型进行横向对比得到更优的预测效果。通过调节训练次数，隐藏层层数等模型参数可以更有效的减小误差。右图为使用了1000个数据作为输入的LSTM模型的预测效果，在后续的开发中，我们也可以通过增加数据量继续减小误差，提升精度。

在模型测试部分，我们引入四种可以衡量模型精确度的度量指标，定量输出模型误差。右图是误差的变化效果，可以看到，随着训练次数增加，各个误差都在不断减小。我们也在代码中加入了Early Stopping模块以监测损失函数（loss），如果在一定数量的训练次数内损失没有显著改善（即没有下降），则会提前结束训练，以防止模型的过拟合。

在软件开发的界面设计部分，我们主要使用python中的Qt-Designer工具设计可视化界面，我们希望我们的软件是拥有多层界面并能实现不同对应功能的，同时也要做到简洁美观，以便用户轻松上手。

我们希望我们的软件能实现多种功能。用户可以自行输入风机的理论和实际的符合格式的数据，并选择不同类型的模型进行训练。同时，用户可以手动调节模型参数，待训练结束后可以直观看到模型的训练结果和各种误差指标。通过理论和实际值分别拟合风电功率曲线，直观反映风机运行效果和风电出力。最后，我们也需要撰写简介实用的说明文档以帮助用户快速上手操作我们的软件

在之后的开发工作中，我们小组利用训练营学到的git及github知识协同工作，控制版本，提高开发效率

我们基于每个人的长处和能力合理安排任务，稳步推进开发进度，有利于快速达成预期目标。

在此次训练营结题答辩之后，我们还会继续跟进大创项目，并对软件开发有更多的后续设想

由于我们目前收集的只是单风机一周内的数据，对于未来的功率预测没有普遍性意义，因此我们还需收集时间范围更广的数据以建立更合理的预测模型，并在之后的软件开发中添加基于时间序列的预测功能。与此同时，我们现在选用的神经网络模型还可以进行进一步优化，达到更高的预测准确度。最后，对界面的不断优化也会同步推进，以提高软件的可用性。

我的报告到此结束，欢迎老师斧正。