

招商证券人工智能工程营课程报告

**项目名称：**

**基于蚂蚁金融的语义匹配模型**

|  |  |
| --- | --- |
| 院系： | 电气与电子工程学院 |
| 小组成员： |  |
| 熊锦慎 | U202112304 |
| 廖茂帆 | U202112291 |
| 郭庭玮 | U202112283 |
| 张洋 | U202112310 |
| 指导教师： | 郑 玮 |
| 2023年9月2日 | |

课程项目设计报告

1 研究背景和项目目标 (对应课程目标1)

1.1选题依据

*基于蚂蚁金融的语义匹配模型具有很强的现实意义，在人工智能客服等多个领域都有很高的实用价值，我们小组出于兴趣选择这个选题希望能对语义匹配方面的人工智能有更深的了解以及学习更多的人工智能知识。*

1.2业界现状介绍

*语义匹配模型需要解决判断两句话语义是否相同的问题；业界现状是已经有了比较完善的匹配系统，但仍有很高的进步空间，有很多人工智能不能解决的问题还是需要人工客服。*

1.3 本项目的目标

*目标：通过训练模型达到可以大批量地进行语义匹配，并且到达较高的准确率。*

2 项目总体设计 (对应课程目标2，5)

*语义匹配模型分为几个板块：数据加载，数据预处理，模型搭建，模型训练 & 评估，模型预测*

3 项目关键技术(对应课程目标2，3)

*为了实现上述板块的任务，需要用到的模型：ERNIE-Gram和ERNIE 3.0 base*

*编程语言：Python；*

*库：import time*

*import os*

*import numpy as np*

*import json*

*import paddle*

*import paddle.nn.functional as F*

*from paddlenlp.datasets import load\_dataset*

*import paddlenlp*

*import random*

*import paddle*

*from paddle import nn*

*from paddle.nn import functional as F*

4 项目实现(对应课程目标3，5，6)

*通过对模型ERNIE-Gram的多次重复训练达到使预测准确率升高的目的，在多次训练之后得到best modle，在最高的准确率处达到收敛。*

5 项目测试(对应课程目标3，4)

*基本实现了语义匹配模型的训练，准确率达到74％左右*

6 项目管理(对应课程目标5，6)

6.1 团队人员组成

廖茂帆，张洋，郭庭玮，熊锦慎

6.2 任务分工

*廖茂帆：模型训练以及汇报展示*

*熊锦慎：资料搜集以及模型搭建*

*郭庭玮：数据预处理板块*

*张洋：模型评估，模型预测*

7. 总结与反思(对应课程目标4，6，7)

*通过本次学习我们对人工智能有了更深刻的了解，虽然没有太深入的学习和挖掘，但在我们心中种下了一颗种子，让我们对人工智能有了更深地兴趣。*

**评分页**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 成员姓名 |  |  |  |  |  |
| 预习报告  16%（百分制） |  |  |  |  |  |
| 平时成绩  20%（百分制） |  |  |  |  |  |
| 课程设计与答辩  64%（百分制） |  |  |  |  |  |
| 合计 |  |  |  |  |  |

**指导教师评语**

|  |  |
| --- | --- |
| 学生1 |  |
| 学生2 |  |
| 学生3 |  |
| 学生4 |  |
| 学生5 |  |
| 教师 | 郑 玮 |

2023年9月16日