ChatGot-基于语音识别和语义理解的智能口语对话平台

摘要

随着全球化进程的不断加快，外语学习的重要性日益凸显。然而，许多学习者在突破口语障碍和与他人实际交流的过程中面临困境。团队基于目前的英语学习困境研究了一款名为“ChatGot 柴狗”的网页平台，致力于帮助用户锻炼英语口语能力。提供智能口语对话和真人匹配两种模式，并辅以交流社区和其他简单功能，旨在为用户提供便捷、高效的英语学习体验。

本项目旨在为英语学习者提供一个实时对话和在线交流的平台。通过智能对话功能，用户可以在不同场景下与AI进行实时对话，提升口语能力和实际场景下的交际能力。同时提供用户匹配功能，使用户可以与其他用户进行自由的英语对话，并通过交流过程中的词条和提示等辅助工具帮助用户进行交流和练习。此外，平台还拥有社交系统，用户可以添加好友、发表动态以及与好友进行聊天，建立一个英语学习的社区环境。

项目关注于实时对话、用户匹配和社交系统等主要功能，以及辅助功能如每日签到、佳句欣赏、单词记忆、新闻阅读和听力音频等，旨在提升用户的英语学习体验和口语水平。

网页前端的设计贴合项目定位，大量的“柴狗”元素除了招人喜欢， 较易让消费者接触之外，在一定程度上有幽默的效果，动物元素能让消费者在轻松的环境中更能接受信息，拉近了与消费者的距离。吸引更多用户的尝试和使用。作为日常生活中代表欢快活泼热情的光辉色彩，暖色系中最温暖的的橙黄色，是一种富足、快乐和幸福的颜色，在现代社会上往往作为标志色和宣传色，属于心理色性，可视性能优良。

通过以上功能的开发和实现，"ChatGot 柴狗"网站旨在帮助英语学习者提升口语水平，增加与他人交流的信心，并为用户提供一个互动的学习平台。同时，该小程序还鼓励用户分享学习心得和寻找语伴，促进用户之间的相互学习和交流。

关键词：英语口语 AI实时对话 社交功能

**选题背景**

“哑巴英语”现象指的是学生虽能阅读英语材料、理解英语听力并进行书面表达，但面对口语交流时却无法流利开口的困境。产生这种现象的原因多种多样，可以归纳为主观和客观两大类。从主观角度看，影响因素包括学习者的口语学习动力、心理障碍等个人因素。而客观层面，则涉及语言知识掌握不足、学习策略缺陷以及教学环境和政策等多方面因素。

《教育信息十年发展规划》（2010-2020年）中强调了数字化教育资源的重要性，推动了从传统教学模式向更互动、多元的教学方式的转变。伴随大数据和人工智能技术的发展，外语教学正在从单向独白式教学转型为更为丰富的多元交互式教学。在这种背景下，许多研究提出利用信息技术手段来改善“哑巴英语”，其中AI技术因其能够针对每位学生提供定制化学习方案而成为破解此难题的关键。

AI技术的引入不仅能缓解传统课堂教学的局限性，还能通过个性化的教学内容和方法来提升学生的学习体验。AI技术特别适合处理语料库的深度和广度问题，它能够提供丰富多样的语言资源，从而使学生在与AI的互动中接触到更广泛的知识和语料。此外，AI技术还能提供地道、准确的语言输入，帮助学生在口语学习中获得更好的支持。

为应对这一挑战，我们设计了“ChatGot”智能英语口语对话平台。该平台利用人工智能技术，模拟真实语境中的对话，帮助学生在各种场景下练习口语，从而提高他们的语言实际应用能力。ChatGot不仅致力于帮助学生从理论知识转向实际应用，还力求让学生在学习过程中体验到乐趣，激发他们的学习热情，最终彻底摆脱“哑巴英语”的困扰。

**实施进程**

立项之前，小组已经对项目应用背景、当前英语学习者面临的一些痛点和需求、以及中国英语教学背景下学生的各项英语能力等进行过资料查阅和调查了解，分析和构建了用户画像。通过查阅文献的方式，对项目的可行性进行了探究；根据设想的内容，制作了网站的界面demo；基于调查结果和项目前景，撰写了项目计划书，详细对项目内容进行了介绍和说明。

项目在2022年十月份与十一月：

1、完成网站后端搭建，实现除了AI实时对话外的绝大多数功能，例如实现用户实时自由匹配，支持用户线上实时交流，实现社交、好友功能等等。

十二月到2023年三月中期：

1、建立初步的智能对话AI，使其可以处理音频数据，具备较高的音频转文本精确度；具备分析文本数据，并能从中提取用户信息，理解用户意图，再基于当前场景的知识源，实现对用户对话的准确答复。

2、完善网站，在美化，排版等多方面进行优化，并发布测试版网站，邀请一些同学对网站的设计提出意见，并按反馈意见进行相应的修改等等，为进行中期答辩做准备。

2023年三月中期答辩后到2023年七月：

1、扩展各个场景下的AI知识源，优化AI对话实际表现；引入智能评价机制，对用户在人机对话中的发言进行打分，评价等，并分析、刻画用户特征；使用推荐算法，根据用户特征优化用户实时匹配功能，使其更可能匹配到水平，偏好与自己相同的用户。

2、针对AI发音存在的问题进行不断修正，尽可能使其发音标准、地道，以免误导学习者。

2023年八月至十一月：

1、对项目的各个方面进行完善和优化，购买了专属的域名并成功将其部署在网站上，使之成为实际运行的平台，准备结项答辩。

2、改进了对话AI，使用了大语言模型来替代之前采取的RASA框架，并对模型进行RLHF微调。

3、为了增强评分算法的表现，引入了评分API，替代成员自制的评分算法。

**成果内容**

主要功能：

1. 实时对话：实现了用户可以和我们的 AI 在特定场景下交流，可以调节语速和音色，使得 AI与使用者在该实际场景下实现交流对话，引导用户实时对话，提升用户的口语能力和实际场景下交际的听力能力。在中期改进之后实现了“实时”效果，给予用户更快速的反馈。

2. 用户匹配：对于希望进一步提高英语交际能力，不满足与 AI 进行特定场景的对话，希望与真人进行自由英语对话或者交友的用户，进入模块后，随机一个房间号并匹配用户，匹配成功后进入房间 （新页面 同时开启通话功能，用户与用户之间可以聊天。交流过程中会有词条和提 示等，协助用户对多个话题进行交流，减少冷场和尴尬的可能，用户也可以自己选择和更新这些 tips或者忽视词条和提示，与对方自由使用英语交流。除此之外，该功能为用户提供了备忘录，用户可以在这里做随手记录。

3. 社交系统：好友交际功能，个人可以发表动态，也可以与好友聊天。同时实现了英语社区环境，在这里可以分享英语的心得和方法，同时也可以公开寻找想要的英语口语练习伙伴。ChatGot 不只局限于人工智能英语对话服务，更是希望通过真人对话的练习模式提高用户的口语水平。

4. 评分功能：我们接入了成熟的API对交流期间用户的发音和回答内容进行评分并且提供建议，让用户知道自己存在的问题，并针对性地进行练习提升。

辅助功能：

平台的辅助功能旨在全面提升用户的英语学习效果，覆盖听力、口语、阅读和写作等多个方面，助力用户提高英语综合能力，丰富学习体验。

1. 每日签到：用户可以在这里进行打卡，方便用户记录每天的学习进展。

2. 佳句欣赏：提供优美句子供用户学习，有助于积累词汇和语感。

3. 单词记忆：这一功能专注于词汇积累，帮助用户加强记忆，对提高英语口语水平尤其有益。

4. 新闻阅读：该功能提供具有时效性的新闻文章，使得用户可以练习英语阅读 以及获取新闻信息等。

5. 听力音频：该功能提供听力使用的音频，用户能够练习听力理解和模仿标准发音。

前端：

网站的前端设计紧密契合了该项目的核心定位，采用了大量“柴犬”元素，不仅深受用户喜爱，也易于吸引用户参与。这些元素在轻松愉快的氛围中带来幽默效果，有效拉近了用户与产品之间的距离。暖色系中的橙黄色调，作为一种象征富足、快乐和幸福的色彩，增强了网页的吸引力和视觉效果。网页的布局设计考虑周全，将主要功能、辅助功能与社区功能进行有效分隔，提升了用户的浏览效率和体验。动态动画和备忘录功能进一步优化了用户界面，缓解了用户的使用压力。这样的前端设计不仅提升了视觉体验，更成为了产品与用户之间沟通的关键桥梁。

**项目创新点**

1、智能互动与真人匹配：平台利用自然语言处理技术打造了AI问答通道和真人匹配通道，为用户提供新颖的练习模式，结合了学习的深度和游戏的趣味性。同时，与市面上流通的功能单一的学习软件不同，ChatGot 平台根据用户意愿匹配其他用户线上交流的功能在解决部分人们由于找不到学伴而缺少练习机会的难题的同时，增加了学习平台的社交属性，满足当代人们的社交需求。

2、情景模拟交流：平台创建了不同生活情景作为口语对话练习背景，使用户能够根据个人需求选择相应场景，实现与 AI 的实时模拟交流，口语练习更加多样化、生活化，更具科技感。

3、智能评分系统：与人工模拟对话相配套的智能评分机制可以为用户提供即时准确的口语表现反馈，让用户体验成长与进步的快乐，激发学习热情的同时还可以定期复盘学习收获。

4、社区互动平台：构建了一个内容丰富的社区交流环境，包括个人主页和社区专区，让用户能够分享学习心得，相互学习，提高口语学习的互动性和动力。

5、词条推送与实战练习：平台具备智能词条推送功能，结合实际话题进行口语实战演练，增强了用户的实用语言能力。

6、亲和力设计：作为潜在用户多为学生党的平台，ChatGot 围绕大众所喜爱的“柴犬”主题开发设计，拉近与用户之间的距离，使用感更加轻松愉快。在视觉上不管是前端的动物元素、主题颜色以及排版布局等潜移默化的影响用户的学习情绪和积极性，创造了一个友好且轻松愉悦的学习氛围，帮助用户改善原本乏味的学习进程，促进用户的积极参与和持续使用。

7、市场独创性：通过专利检索和市场调研，ChatGot证明了其产品设计的独创性，市场上未发现具有相同创新功能的同类产品。

8、辅助学习功能：除了基础学习功能外，平台还提供了包含听说读写的全方位辅助工具，多元化的辅助功能可以全面提升用户的语言水平，帮助用户提高学习兴趣和知识储备，给予用户一站式全流程服务和体验。

**成果应用情况：**

目前本次项目并没有申请专利、发布论文或者出版专著。但是我们已经购买域名并部署到网站上进行使用。如果网站被更多用户所发掘并喜爱，“柴狗”就可以帮助到更多想要解决“哑巴口语”问题的人练习口语，也可以成为英语学习爱好者的专属社区，具有很强的教育领域意义。从商业角度出发，假设“柴狗”拥有了一部分用户并愿意为了某些特定功能买单，我们可以优化算法，优化界面，在不断吸收客户的同时开发收费的项目，让用户愿意为在线口语对话平台买单，从而实现收益。

**研究心得：**

1. 框架设计

学会了一门新的框架 django ，对前后端的制作更加熟练。遇到的问题是在导入模块，以及对后端发送请求，收到请求的处理一开始不是很熟练，查阅了相关的文档之后解决了很多。

2. 评价算法

对四种算法的原理和运作有了新的认识，我们也学习了其他大厂研发的评分算法，阅读了相关的论文，学习到了很多的知识。在开发中主要遇到的问题在于评价体系中参考标准的选择和获取、如何评价用户的音素以及如何不通过语音转文字来对用户的输入进行评估以提高评价的精确度。在查阅资料和自主实施后解决了很多问题，包括语法错误、功能偏差等多个方面。我们首先通过自行标注训练试图提高评分的准确性，但这样的做法并不适合小团队开发，所以后期我们选择了接入其他成熟的API来进行评分。

3. 语言模型

作为一个智能对话模型，训练一个表现合格、反应迅速、结果可控的语言模型是重中之重。最初由于团队成员深度学习功底不足，我们选择使用RASA3框架，根据开源对话语料以及团队成员自己构造的语料，训练了第一版检索式对话AI，doggo-rasa。它的好处在于反应十分迅速，在智能对话流程中，AI生成回复耗时非常少，由于不需要较高的算例，它还可以直接部署在本机使用；同时，由于模型是检索式的，它的回复结果高度可控，不会产生不健康、不合适的回复。但是在后来的测试中，我们团队认为doggo-rasa作为一个对话AI，还不够智能，回复模式单一，缺乏灵活性。因此在中后期，团队引入了开源大模型chatGLM作为对话AI，并在chatGLM3发布时，及时更新模型版本。chatGLM3-6B的部署和回复需要一定算力，因此我们将他单独部署在了GPU云服务器上，通过网络接口与之交互。我们使用基于人类反馈的指令微调策略，在6个特定的对话场景下，对模型进行了微调，使得模型在我们设定的6个对话情景下有着更加优良的表现，得到了我们现在正在使用的doggo-chatglm。它有着相当优秀的表现，在多数情况下都可以与用户展开顺畅的对话，且允许用户与其自由对话，而不只是限制在固定的场景。

4. 团队整体

初期遇到了很多大大小小的问题，甚至在用 github push 代码的时候都会出错。到后期初步开始写前端和后端的时候更是会遇到很多 bug，比如包的版本不匹配，前一天还能正常运行的代码放到第二天就不能跑了，同样的代码在不同成员的电脑上运行会报不一样的错误，中途遇到全新的从未学习过的新技术和框架需要付出额外的学习成本等等。对此我们都积极上网查找资料，试图可以站在巨人的肩膀上运用网络上前人的经验积累来解决自己的问题，搜索答案的渠道有很多，csdn、stackoverflow、github，甚至 chatgpt 都是我们解决问题的主要途径之一。在中期对项目方向产生迷茫时，我们也积极向导师寻求帮助，结合导师的意见将项目的重心稍作调整。在项目实施的过程中，我们解决问题的思路更加开阔了，懂得从多种渠道寻求帮助，在遇到问题时的心态更加成熟，应对各种各样的 bug 也更加熟练。创新创业比赛是团队的比赛，队员之间的合作和沟通尤为重要，我们团队的成员一直秉承着“有问题及时沟通”的原则，遇到过类似问题的成员可以帮助其他成员解决问题，以此节省时间，解决全组的学习成本，通过这次项目合作经历，我们也更加懂得了团队合作的重要性。

除此之外我们每个人在做自己负责的内容时也遇到了各自不一样的问题，有自己的感受和学习经验获得。本次双创项目在多个维度上对我们都有很大的提升，比如查询收集资料、前后端的制作、处理 bug 等在很大程度上提高了我们的信息检索、编程能力也增强了心理承受能力和充分的耐心。除了项目本身，在团队协作方面也给每个人带来了很多感悟，如何分配工作、和组员协调、如何调节组内氛围等也是一个方面的提升，在这个小组里，我们不仅可以讨论自己遇到的问题寻求组员的帮助，也可以询问和讨论学业上的困惑、分享自己的新的体会、帮助他人缓解焦虑情绪等情感辅助。在项目推进的过程中学会欣赏他人的能力、发现自己的不足并及时修补、主动向他人学习和参与讨论。