

学习计划

我叫纳吉 (Rahaman Nagiur)，对人机交互和技术之间的错综复杂关系有着浓厚的兴趣，这驱动我踏上计算机科学的旅程。人工智能不仅理解，还能复制人类语言的细微之处，特别是情感和表达特征，这为我们如何与技术互动和获取信息带来了巨大的潜力。这种潜力在多媒体领域尤其引人注目，因为逼真且引人入胜的数字语音对于吸引受众和促进有效沟通至关重要。这种愿望促使我追求计算机科学与技术硕士学位，该项目以其在文本转语音(TTS)、智能语音分析、语音增强、自动语音识别(ASR)、多媒体分析、声学场景分析和人机交互等领域的创新研究而闻名。此外，该项目重点关注先进的深度学习技术，包括迁移学习、对抗学习、半监督学习、无监督学习、主动学习和自动机器学习(AutoML)，这提供了发展技能以推进这一快速发展的领域的机会。我的最终目标是作为一名研究科学家为多媒体行业做出贡献，专注于创建最先进的多语言TTS系统，并最终在该领域创办自己的初创公司。

我的本科学习，计算机科学与技术专业，在兰州理工大学为我奠定了扎实的理论和实践基础。机器学习、深度学习、自然语言处理、大数据技术、数据结构与算法、SQL和软件工程等核心课程，让我掌握了解决复杂计算问题的基本原理。我的学术表现一直非常出色，并荣获兰州理工大学两项优秀学生奖。这些奖项不仅认可我的优异成绩，更彰显了我对学习的投入和将理论知识应用于实际问题解决的能力。

我的本科毕业论文深入探讨了情感表达文本转语音合成的复杂性。该研究项目要求我应用深度学习模型，特别是Transformer模型，来合成语音中的各种情感特征。项目涵盖了数据收集、预处理、模型训练和评估。至关重要的是，该论文经历磨练了我的研究方法，从提出研究问题到严格评估结果。我熟练掌握的编程语言，例如Python和JavaScript，通过独立项目以及与教授的合作得到了进一步的强化，从而提升了我的实际编程技能。

我的长期目标是利用我的技术专长，为多媒体行业做出重大贡献。我希望自己成为一名研究科学家，积极参与开发创新的TTS（文本转语音）解决方案，以满足全球受众的多样化沟通需求。最终目标是创办一家多语种TTS初创公司，使不同社群能够通过细致入微且逼真的数字声音有效地连接和沟通。本硕士项目将对这一雄心勃勃的目标至关重要，它将使我能够获得将研究转化为商业化产品的必要技能和知识。我的目标不仅仅是构建一个系统，而是构建一个真正引人入胜、富有同理心且切合用户需求的系统。

除了学术追求，我还积极寻求机会提升我的实践技能和经验。我的研究论文“微调预训练语言模型进行语法可接受性、纠错、情感分析和情绪检测”，已发表在《国际高级工程技术研究杂志》上。该论文突显了我进行独立研究、设计实验、分析数据以及有效地将研究成果传播给更广泛的学术界的才能。我在兰州理工大学领导的AI学习小组的经历，包括项目管理、指导其他学生以及有效沟通，为我在研究环境中必不可少的团队合作技能提供了宝贵的经验。这些经历也让我深刻理解了合作的重要性，以及通过技术解决现实问题的意义。

我的研究兴趣在于开发一个超越现有局限性的强大多语言TTS框架。我特别关注以下几个具体的研究领域：1) 探索TTS中合成细微情感表达的新方法；2) 开发准确捕捉和再现语言特定韵律特征的方法；3) 发展用于创建可定制语音的说话人属性控制技术。这些领域需要深入理解先进的深度学习技术，例如迁移学习和对抗学习，以及先进的声学建模。此外，我希望探索如何应用半监督或无监督学习技术，以提高TTS系统的效率和泛化能力，同时减少对大量标记数据的需求。我的目标是开发一个在多种语言和文化中展现出卓越性能的框架，为多语言TTS系统树立新的标准。

硕士毕业后，我渴望加入专注于语音技术的科研实验室或初创公司。我的目标是利用我的知识和技能，进一步开发强大的TTS模型，并为该领域的尖端研究做出贡献。我希望积极参与基于先进TTS系统的商业化产品的开发，解决实际问题，并对多媒体内容的可访问性和易用性产生切实的影响。

我相信贵校计算机科学与技术专业的硕士项目能够提供理想的环境，让我进一步提升技能，并在语音技术领域做出重大贡献。我的研究兴趣、职业抱负以及相关经验与该项目专注于尖端研究、先进深度学习技术和强大师资力量高度契合。我渴望加入充满活力的研究群体，为TTS和智能语音分析的持续进步做出贡献，最终影响人类与计算机交互和多媒体可及性的未来。我相信我的投入，加上贵校提供的资源和指导，将为该领域带来有影响力的创新和显著的进步。