

《自然语言处理》课程教学大纲

课程英文名	Natural Language Processing				
课程代码	B050165s	课程类别	专业课	课程性质	专业选修
学 分	2		总学时数	32	
开课学院	计算机学院		开课基层教学组织	智能信息处理课程组	
面向专业	计算机科学与技术、智能计算与数据科学(计算机科学与技术)、计算机科学英才班(计算机科学与技术)、软件工程		开课学期	第 5 学期	

注：课程类别是指学科基础课/专业课/实践环节/通识公共课/公共基础课/其他；

课程性质是指通识必修/通识选修/学科必修/专业必修/专业选修/实践必修/实践选修。

一、 课程目标

自然语言处理 (Natural language processing, NLP) 或计算语言学是信息时代最重要的技术之一。NLP 的应用无处不在，因为人们几乎用语言交流一切：网络搜索、广告、电子邮件、客户服务、语言翻译、虚拟代理、医疗报告等。本课程为自然语言处理的基础教程，从计算的角度来研究人类语言。本课程的主要重点是理解课程大纲中列出的各种 NLP 任务、有效解决这些问题的算法以及评估其表现的方法。将重点关注统计和神经网络学习算法，这些算法在文本语料库上进行训练，以自动获取执行任务所需的知识。课堂讲授将讨论一般问题以及目前的抽象算法，并提供一些算法的实现版本，作为课程项目的一部分进行扩展和实验。

课程的主要知识单元有：系统性介绍自然语言处理常用的理论知识，包括自然语言处理任务限制、技术范畴、语料库、中文自动分词、数据预处理、自然语言处理工具、数据可视化等及相关算法，目的是使学生学会自然语言处理相关的基础知识，以便为应用涉及的任务分析、技术选型奠定基础。同时介绍自然语言处理的命名实体识别、文本信息抽取、文本分类、问答与摘要、机器翻译、语义计算、知识图谱、情感分析等技术领域与重要应用，学习自然语言处理常见应用任务的内涵和解决这些任务的常用方法。

通过理论教学和实践活动，达到以下课程目标：

课程目标 1：能够理解自然语言处理的基本概念，包括研究内容和应用领域。能够阅读有关自然语言处理的专业论文，具备阅读专业论文的能力

课程目标 2：掌握基于概率统计模型的自然语言处理方法，能够利用现有自然语言处理工具进行自然语言相关论题的理论与实验研究

课程目标 3: 能够利用本课程及论文阅读所掌握的知识, 选定某个主题, 开发实现一个自然语言处理领域的实验系统

课程目标 4: 具备基本的科学素养和创新意识, 在课程实践的过程中, 及时了解自然语言处理的国内外新技术和发展趋势, 及时掌握国家相关方面的科技战略需求, 树立强烈的爱国主义使命感与责任心, 同时能客观分析人工智能, 特别是自然语言处理技术发展中涉及到的伦理和社会问题。

二、 课程目标与毕业要求对应关系

根据计算机科学与技术专业培养方案中毕业要求指标点的分解、课程与毕业要求的对应关系表, 《自然语言处理》支撑毕业要求 1 的指标点 1-3 和 1-4, 毕业要求 2 的指标点 2-2, 毕业要求 4 的指标点 4-3, 毕业要求 8 的指标点 8-1, 毕业要求 10 的指标点 10-1, 课程目标与毕业要求、指标点的对应关系如表 1 所示。

表 1 课程目标与计算机科学与技术专业毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1: 工程知识: 掌握数学、自然科学、工程基础、计算机专业领域的知识, 并能应用于计算机领域复杂工程问题的解决方案中。	1-3 能够运用计算机专业知识, 对计算机领域复杂工程问题解决方案进行分析与优化。	目标 2: 0.4 目标 3: 0.6
	1-4 掌握计算机系统、物联网、人工智能、大数据、网络安全等某个专业领域的知识, 并用于解决计算机领域的复杂工程问题。	目标 2: 0.4 目标 3: 0.6
毕业要求 2: 问题分析: 能够应用数学、自然科学和工程科学的基本原理, 以及科学思维方法, 对计算机领域的复杂工程问题进行识别、表达和分析, 并通过文献查阅与研究获得有效结论。	2-2 具备文献检索能力, 能够对计算机相关领域的复杂工程问题进行调查、研究分析。	目标 1: 0.6 目标 3: 0.4
毕业要求 4: 研究: 能够基于包括计算学科在内的科学原理, 采用科学方法研究计算机领域的复杂工程问题, 包括设计实验、分析与解释数据、并通过信息综合得到合理有效的结论。	4-3 能够收集、分析与解释数据, 并通过信息综合得到合理有效的结论。	目标 1: 0.2 目标 2: 0.3 目标 3: 0.5
毕业要求 8: 职业规范: 具有人文素养和社会责任感, 能够在计算机工程实践中理解伦理道德、遵守职业规范、履行社会责任。	8-1 具有人文社会科学素养、社会责任感。	目标 1: 0.2 目标 4: 0.8
毕业要求 10: 沟通: 具备一定的国际视野和跨文化沟通能力, 能够就专业领域的复杂工程问题与业界同行及社会公众有效沟通, 包括文字表达和语言交流。	10-1 能够就计算机复杂工程问题撰写报告和设计文稿、陈述发言、清晰表达或回应指令。	目标 2: 0.4 目标 3: 0.6

根据软件工程专业培养方案中毕业要求指标点的分解、课程与毕业要求的对应关系表, 《自然语言处理》支撑毕业要求 1 的指标点 1-3 和 1-4, 毕业要求 2 的指标点 2-2, 毕业要求 4 的指标点 4-

3, 毕业要求 8 的指标点 8-1, 毕业要求 10 的指标点 10-1, 课程目标与毕业要求、指标点的对应关系如表 2 所示。

表 2 课程目标与软件工程专业毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1: 工程知识: 掌握数学、自然科学、工程基础、软件工程专业领域的知识, 并能应用于软件工程领域复杂工程问题的解决方案中。	1-3 能够运用软件工程专业知识, 对软件工程领域复杂工程问题解决方案进行分析与优化。	目标 2: 0.4 目标 3: 0.6
	1-4 掌握某个专业领域知识, 并用于解决软件工程领域复杂工程问题。	目标 2: 0.4 目标 3: 0.6
毕业要求 2: 问题分析: 能够应用数学、自然科学、工程科学的基本原理和科学思维方法, 对软件工程领域复杂工程问题进行识别、表达和分析, 并通过文献查阅与研究获得有效结论。	2-2 具备文献检索能力, 能够对软件工程领域复杂工程问题进行调查、研究分析。	目标 1: 0.6 目标 3: 0.4
毕业要求 4: 研究: 能够基于包括计算学科在内的科学原理, 采用科学方法研究软件工程领域复杂工程问题, 包括设计实验、分析与解释数据、并通过信息综合得到合理有效的结论。	4-3 能够收集、分析与解释数据, 并通过信息综合得到合理有效的结论。	目标 1: 0.2 目标 2: 0.3 目标 3: 0.5
毕业要求 8: 职业规范: 具有人文素养和社会责任感, 能够在计算机工程实践中理解伦理道德、遵守职业规范、履行社会责任。	8-1 具有人文社会科学素养、社会责任感。	目标 1: 0.2 目标 4: 0.8
毕业要求 10: 沟通: 具备一定的国际视野和跨文化沟通能力, 能够就专业领域的复杂工程问题与业界同行及社会公众有效沟通, 包括文字表达和语言交流。	10-1 能够就软件工程领域复杂工程问题撰写报告和设计文稿、陈述发言、清晰表达或回应指令。	目标 2: 0.4 目标 3: 0.6

根据计算机科学英才班(计算机科学与技术)专业培养方案中毕业要求指标点的分解、课程与毕业要求的对应关系表,《自然语言处理》支撑毕业要求 1 的指标点 1-3 和 1-4, 毕业要求 2 的指标点 2-2, 毕业要求 4 的指标点 4-3, 毕业要求 8 的指标点 8-1, 毕业要求 10 的指标点 10-1, 课程目标与毕业要求、指标点的对应关系如表 3 所示。

表 3 课程目标与计算机科学英才班(计算机科学与技术)专业毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1: 工程与科学知识: 能够掌握数学、自然科学、工程基础、计算机科学技术理论知识, 并应用在计算机相关领域的复杂工程问题和基础科学问题的解决方案中。	1-3 能够运用计算机专业知识, 对复杂工程问题解决方案进行分析与优化。	目标 2: 0.4 目标 3: 0.6
	1-4 掌握计算机系统、人工智能、大数据、网络安全等某个专业领域的知识, 并用于解决计算机相关领域的复杂工程问题。	目标 2: 0.4 目标 3: 0.6