

优 必 选

UBTECH



智能语音信号处理及应用



项目 5

智能客服之对话机器人

任务一 对话机器人的关键技术：语音识别

CONTENT 目 录

01 课程导入

02 语音识别的概念

03 语音识别的原理

04 语音识别的应用

05 语音识别的应用

06 总结



课程导入

“嘿 Siri, 用免提给老妈打电话”



唤醒语音助手，可以直接实现拨打电话！多么便捷的操作！如何实现的呢？

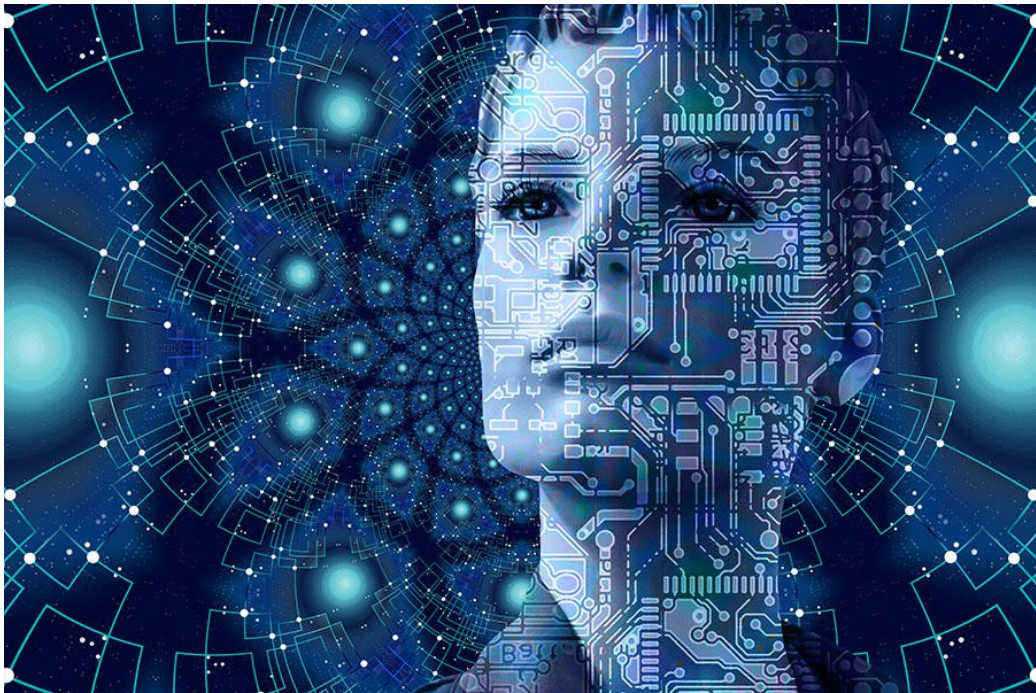


语音识别是将语音输入至智能设备，通过其对语音信号进行识别和理解为文本或命令的技术。



语音识别的概念

语音识别技术，其涉及数字信号处理、人工智能、语言学、数理统计学、声学、情感学及心理学等学科。应用范围，主要包括工业、军事、交通和民用等领域。



语音识别的原理

模仿或代替人耳的听觉，本质上是一种模式识别。

原理为模式匹配，通过将输入的未知语音模式与已有的语音模式进行对照比较，匹配到的最佳模式为最终识别结果。



语音识别的原理

语音识别的步骤：训练和识别。

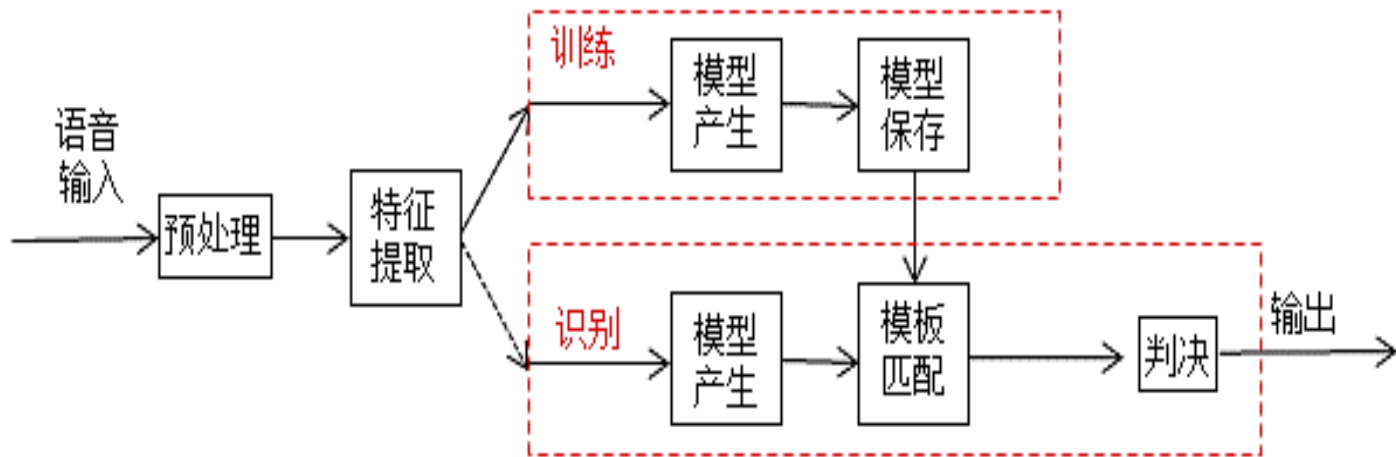
第一步是学习（训练）的过程，根据识别系统的类型进行识别方法的选择和语音特征参数的分析后将其存储；

第二步对输入语音进行识别。



语音识别的原理

特征提取，该过程中需要对特征参数进行选择。选取合适的特征进行提取完毕后，测试前进行模型的训练，保存训练好的模型；最后在识别时根据失真判决准则进行判别。



语音识别的原理

特征提取的方法：

线性预测系数（LPC）、感知加权预测系数（PLP）、短时平均过零率、线性预测倒谱系数（LPCC）、自相关函数、梅尔倒谱系数（MFCC）、小波变换系数、经验模态分解系数（EMD）、伽马通滤波器系数（GFCC）等。



语音识别的分类

1. 按识别单位划分，从易到难为；
2. 按识别词汇量划分；
3. 按讲话人的范围划分；
4. 按识别方法划分；
5. 按识别环境划分；
6. 按传输系统划分；
7. 按说话人的类型划分。



语音识别的分类



按识别单位划分，从易到难为：

- （ 1 ）孤立单词语音识别。所识别单词与单词之间有间歇。
- （ 2 ）选词语音识别。识别连续语音中的某个或某几个单词。
- （ 3 ）连续语音识别。识别连续的单词，单词间不停顿。
- （ 4 ）语音理解。不仅对语音中的单词进行识别，还根据语言学知识进行含义的推理。



语音识别的分类



按识别词汇量划分，该方法识别量通过词汇表中包含的词条来计算，且识别率将随着单词量增多减小。

- (1) 小词汇。识别 10-50 个孤立词。
- (2) 中词汇。识别 50-200 个孤立词。
- (3) 大词汇。识别 200 个以上孤立词。



语音识别的分类

按讲话人的范围划分。

（ 1 ）单个特定讲话人识别。基于用户输入大量发音数据进行训练，属于特定人语音识别；

（ 2 ）多讲话人（有限讲话人）识别。从大量发音数据中学习非特定人的基本特征，寻找分析特征的相似性，作为识别的标准，属于非特定人的语音识别；

（ 3 ）与讲话者无关（无限讲话人）识别。与第二种类似，需要对非特定人的基本特征进行分析作为识别标准，属于非特定人的语音识别。

语音识别的分类

e

按识别方法划分

(1) 模板匹配法。根据测试语音与模板的参数比较结果进行匹配，通过失真测度最小准则进行判决。

(2) 随机模型法。根据隐马尔科夫模型 (HMM) 的概率参数进行似然函数估计，然后进行判决得到识别结果。HMM 的状态函数使得随机模型法能利用语音频谱内在变化和它们的相关性，在识别中应用语言结构的动态特性

(3) 概率语法分析法。利用连续的语言中形式语法约束的知识进行似然函数的估计和判决，适合对大范围较长的连续语言进行识别。

语音识别的分类



按识别环境划分，其中在不同的环境下，主要分为以下两种：

- (1) 隔音计算机房识别；
- (2) 公共场合识别。



按传输系统划分。

- (1) 高质量话筒；
- (2) 电话；
- (3) 近讲话筒。



按说话人的类型划分

。

- (1) 男声；
- (2) 女声；
- (3) 儿童声。

语音识别的应用

人与机器之间不能直接进行交流，因此语音识别技术的出现使得人与机器之间能够进行信息传递，达到简单高效的沟通。

在生活中，最常见的语音识别技术是智能手机中的语音输入法、语音助手、语音检索等。



语音识别的应用



一些智能家居中，机器通过语音识别技术听懂人的意图，进行智能电视、空调、照明系统等控制，智能可穿戴设备、智能车载设备中的语音交互也是语音识别技术。



语音识别的应用

在传统行业之中有大部分应用了语音识别技术，例如医院通过语音进行电子病历的录入，法庭的庭审现场通过语音识别进行文字记录，影视中通过语音识别进行字幕制作，呼叫中心的录音质检，听录速记等。



语音识别的应用



除此之外，语音识别技术还与其他技术相结合被应用在翻译上，利用口语识别技术、机器翻译技术和语音合成技术来实现不同语言之间的无障碍交流，例如自动口语翻译。



语音识别的概念

语音识别的原理

语音识别的分类

语音识别的应用



优 必 选

UBTECH



智能语音信号处理及应用



项目 5

智能客服之对话机器人

任务二 对话机器人的关键技术：自然语言处理

CONTENT 目 录

01 课程导入

02 自然语言处理的概念

03 自然语言处理的原理

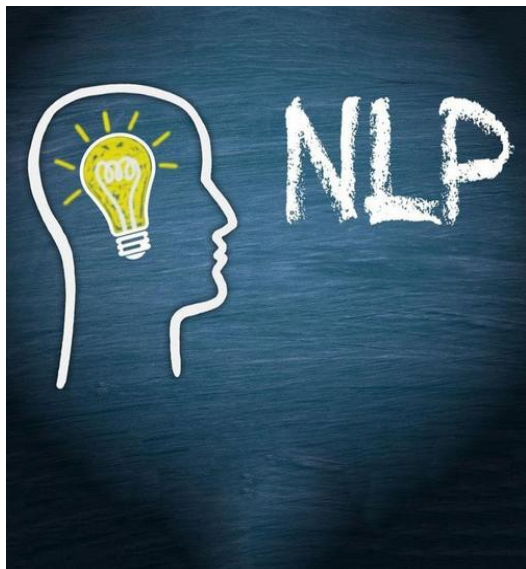
04 自然语言处理技术分类

05 自然语言处理的应用

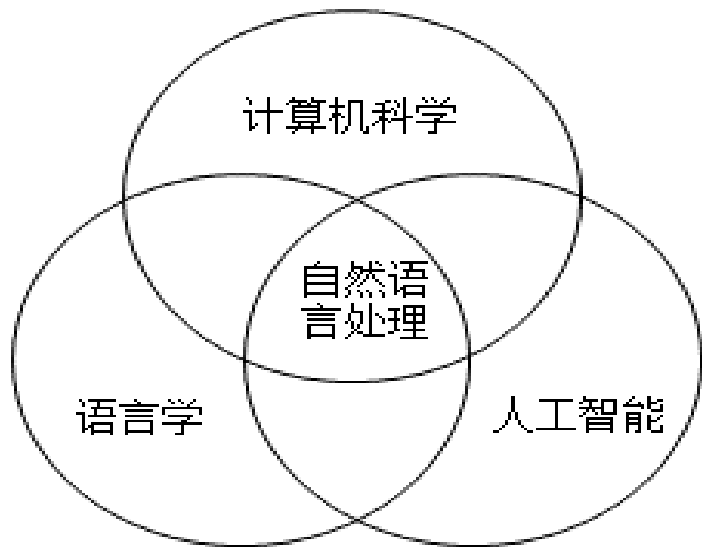
06 总结



课程导入



自然语言处理的概念



该学科主要研究如何通过机器学习技术，让计算机对人类语言进行处理，最终让机器能够对人类语言进行理解。



自然语言处理的原理

处理机制中包括自然语言理解和自然语言生成两个过程，其中前者是指计算机可以理解自然语言文本的意义，后者则是用自然语言文本表达给定的意思。

自然语言处理的具体应用和表现形式有：机器翻译、文本摘要、文本分类、文本校对、信息抽取、语音合成、语音识别等。



自然语言处理的原理

自然语言的理解分析表现为层次化的过程。



自然语言处理的原理

自然语言生成主要可以划分为三个阶段：

1. 文本规划：对结构化数据中基础内容进行规划。
2. 语句规划：从结构化数据中组合语句，表达信息流。
3. 实现：用语法通顺的语句表达文本，NLP 与文本挖掘（或文本分析）之间的不同。



1. 词法分析

包括词性标注和词义标注，其中词性是词汇的基本属性。词义标注关键在于解决一些多义词在具体语境中的意思选择。

标注步骤一般为先确定语境，然后明确语义。



2. 句法分析

断句子的句法结构和组成句子的各成分，以明确它们之间的相互关系。句法分析分为完全句法分析和浅层句法分析。完全句法分析是通过一系列的句法分析过程，确定句子包含的所有句法信息，并确定句子中各成分之间的关系，最终得到一个句子的完整的句法树。浅层句法分析又称部分句法分析或语块分析，它只要求识别出语块，也就是句子中结构相对简单的成分。



3. 语义分析

语义分析是指根据上下文，通过句子的句法结构和句子中每个实词的词义判断该句子的意义，并以某种形式表示，即将自然语言转化为能够被计算机理解的形式语言。句子的分析与处理过程，主要采用“句法语义一体化”的策略，另外还有一些会采用“先句法后语义”的方法。



4. 语用语境和篇章分析

语用是指人在一定的语境下对语言的具体运用，分析并研究语言使用者的用意，它与语境、语言使用者的知识涵养、言语行为、想法和意图均有关，是对自然语言的深层理解。语境分析主要包括情景语境和文化语境两个方面，篇章分析则是不仅要分析单个句子，更要对段落以至整篇文章进行理解和分析。



1. 机器翻译

指利用计算机把一种自然语言转变为另一种自然语言的过程。是实现源语言和目标语言之间的转变。

2. 信息检索

信息检索是指在信息集合中，查找满足用户信息需求的过程和技术。



3. 文本情感分析

情感分析是一种广泛的主观分析，它通过自然语言处理技术对带有情感色彩的主观性文本进行分析。

4. 自动问答

自动问答是指利用计算机自动回答用户所提出的问题以满足用户知识需求的任务。不同于现有的搜索引擎，问答系统是一种高级的信息服务，系统返回用户的是精准的自然语言答案，而不再是基于关键词匹配排序的文本。

5. 自动文摘

通过计算机将给定的一篇文档或多篇文档进行自动分析，提炼，最终总结出一篇长度较短、可读性良好的文本摘要。

6. 信息抽取

信息抽取是指从网页、新闻、论文文献、微博等非结构化或半结构化文本中提取特定的事件或事实信息，并通过信息合并、消除冗余和冲突消解等手段将非结构化文本转换为结构化信息的一项综合技术。

7. 文本挖掘

文本挖掘的准备工作由文本收集、文本分析和特征修剪三个步骤组成，文本挖掘是信息挖掘的一个研究分支，主要用从原本未经处理的文本中提取出未知的知识。

8. 语音识别

语言识别指的是利用语言的统计和语法属性来将不同语言的文本区分出来。语言识别也可以被认为是文本分类的特殊情况。



自然语言处理的应用

1. 知识图谱

主要被用在搜索引擎中丰富搜索结果，并体现搜索结果之间的关联性。知识图谱的划分可以按照应用方式，将其分为语义搜索知识问答、以及基于知识的大数据分析和决策等。

语义搜索利用建立大规模知识库对搜索关键词和文档内容进行语义标注。知识的问答必须要基于知识库，通过对提问句子的语义分析，在将其解析为结构化的询问，在已有的知识库中获取答案。



自然语言处理的应用

2. 机器翻译

机器翻译作为自然语言信息处理的一个分支，能够轻松地将一种语言的文本转换为另一种语言的文本。

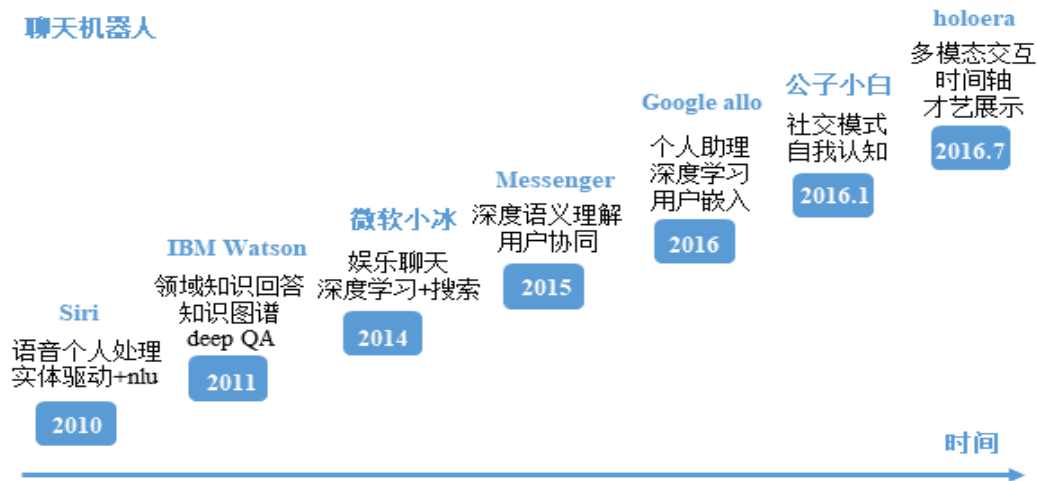


自然语言处理的应用

3. 聊天机器人

聊天机器人是指能通过聊天 app、聊天窗口或语音唤醒 app 进行高效且持续交流的计算机程序，可以解决客户的基本查询，且成本较低。

聊天机器人



自然语言处理的应用

4. 文本分类

文本分类是指将大量的文档根据其内容或者属性不同归类为一个或多个类别的过程。这一技术的关键问题是如何构建一个分类函数或分类模型，将未知文档映射到给定的类别空间。



自然语言处理的应用

5. 搜索引擎

搜索引擎的主要任务是通过用户输入的需求帮助用户找到答案，是连接人与实体世界的服务。最基本的模式是自动化地解析、处理和组织搜集到的内容，并响应用户的搜索请求，找到对应结果返回。

Google
谷歌

YAHOO!
中国雅虎

SOSO 搜搜
腾讯旗下搜索网站

中搜
zhongsou.com

bing
必应 Beta

ASK 爱问

Dogou 狗狗

Sogou 搜狗

Baidu 百度

有道 youdao
网易旗下搜索



自然语言处理的应用

6. 个性化推荐

自然语言处理可以依据大数据和历史行为记录。分析出用户的特征和兴趣爱好，理解用户意图，同时对语言进行匹配计算，实现精准预测符合用户特征和偏好的信息或商品。



自然语言处理的概念

自然语言处理的原理

自然语言处理技术分类

自然语言处理的应用



优 必 选

UBTECH



智能语音信号处理及应用



项目 5

智能客服之对话机器人

任务三 对话机器人的关键技术：语音合成

CONTENT 目 录

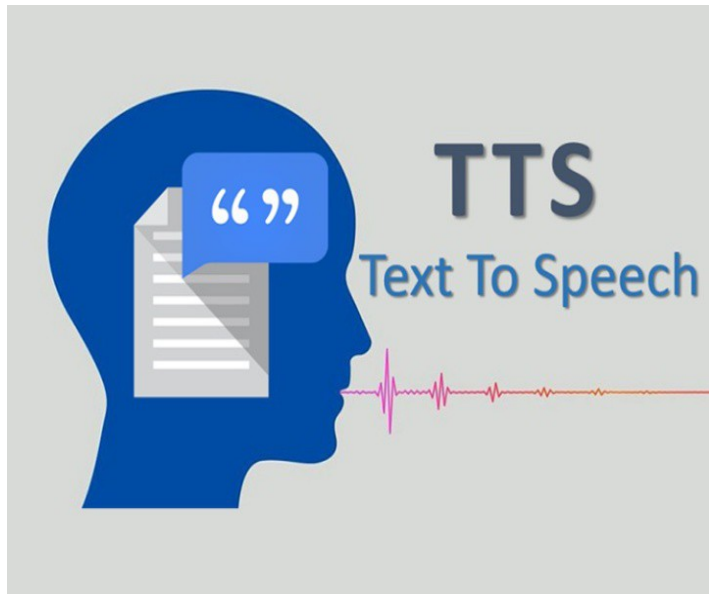
01 课程导入

02 语音合成的概念

03 语音合成技术的应用

04 总结





TTS 技术是什
么呢？



语音合成的概念

语音合成，是通过机械的、电子的方式产生的人造语音技术，它可将任意文字信息实时转化为可以听懂的、标准流畅的语音并朗读出来，让机器像人一样开口说话。

语音合成解决的主要问题就是如何将文字信息转化为可听的声音信息。

声音产生之前会在大脑中产生一段高级的神经活动，即先产生一个说话的意向，然后将意向转换成一系列相关的概念，最后将这些概念组织成语句发音输出。



语音合成的概念

语音合成大致可以分为三类：

1. 文语转换 (text-to-speech) ，即按规则从文本到语音的合成；
2. 概念语音 (concept-to speech) ，即按规则从概念到语音的合成；
3. 意向语音 (intention-to-speech) ，即按规则从意向到语音的合成。



语音合成技术的应用

各种场合的自动报时、自媒体短视频、语音助手、餐饮叫号、电话自动查询服务和文本校对中的语音提示等。



语音合成技术的应用

在电信声讯服务领域中的智能电话查询系统，使用语音合成技术能够实现海量数据和动态查询的需求，可查询一些动态信息。在智能硬件中，能与使用用户进行多语言的对话交流。

将语音合成技术与其他技术相结合。



语音合成的概念

语音合成技术的应用



谢谢大家！

