**Професионална гимназия по електроника и енергетика, гр. Банско**

**България, Банско 2770, ул. "България" №23,**

**тел: 074988402, е-mail: pgeebansko@abv.bg**

|  |
| --- |
|  |

Дипломна работа

Тема:

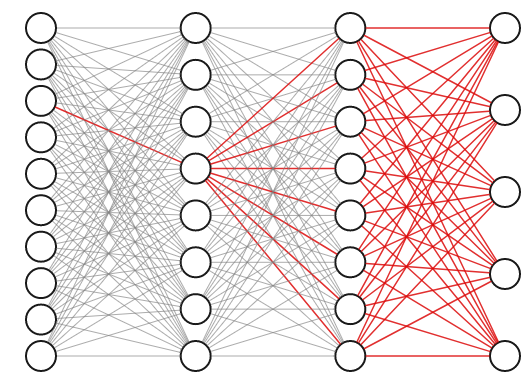
Програма на Python за решаване на квадратно уравнение

Банско 2023

Невронни мрежи

Представете си, че сте планински катерач на върха на планина и е паднала нощ. Трябва да стигнете до базовия си лагер в подножието на планината, но в тъмнината само с малкото си фенерче не можете да видите повече от няколко метра пред себе си. И така, как слизате? Една стратегия е да погледнете във всяка посока, за да видите в коя посока земята се спуска най-много, и след това да пристъпите напред в тази посока. Повторете този процес много пъти и постепенно ще се спускате все по-надолу. Понякога може да заседнете в малка падина или долина, в който случай можете да следвате инерцията си малко по-дълго, за да излезете от нея. Като оставим настрана предупрежденията, тази стратегия в крайна сметка ще ви отведе до дъното на планината. Този сценарий може да изглежда отделен от невронните мрежи, но се оказва добра аналогия за начина, по който са обучени. Всъщност толкова добра, че основната техника за това, градиентно спускане, звучи много като това, което току-що описахме. Спомнете си, че обучението се отнася до определяне на най-добрия набор от тегла за максимизиране на точността на невронната мрежа. Невронните мрежи могат да се използват, без да се знае как точно работи обучението, точно както човек може да работи с фенерче, без да знае как работи електрониката вътре в него. Повечето съвременни библиотеки за машинно обучение са автоматизирали значително процеса на обучение. Поради тези неща и тази тема, която е по-строга от математическа гледна точка, може да се изкушите да я оставите настрана и да се втурнете към приложения на невронни мрежи. Освен това способността за обучение на големи невронни мрежи ни убягваше в продължение на много години и едва наскоро стана осъществима. Невронна мрежа със скрити слоеве са силно взаимозависими, за да видите защо, помислете за подчертаната връзка в първия слой на трислойната мрежа по-долу. Ако променим леко теглото на тази връзка, това ще повлияе не само на неврона, към който се разпространява директно, но и на всички неврони в следващите два слоя, и по този начин ще повлияе на всички резултати.

фиг. 1



И поради сложността за обучение на невронни мрежи ние ще използваме готови обучени невронни мрежи ще разгледаме кои са най-популярните за разпознаване на реч.

**Как работи разпознаването на реч – Общ преглед**

Първият компонент на разпознаването на реч е, разбира се, речта. Речта трябва да се преобразува от физически звук в електрически сигнал с микрофон и след това в цифрови данни с аналогово-цифров преобразувател. Веднъж дигитализирани, няколко модела могат да се използват за транскрибиране на аудиото в текст.

За да се декодира речта в текст, групи от вектори се съпоставят с една или повече фонеми - основна единица на речта. Това изчисление изисква обучение, тъй като звукът на една фонема варира от говорещ до говорещ и дори варира от едно изказване на друго от един и същи говорещ. След това се прилага специален алгоритъм за определяне на най-вероятната дума (или думи), които произвеждат дадената последователност от фонеми. За да се декодира речта в текст, групи от вектори се съпоставят с една или повече фонеми - основна единица на речта. Това изчисление изисква обучение, тъй като звукът на една фонема варира от говорещ до говорещ и дори варира от едно изказване на друго от един и същи говорещ. След това се прилага специален алгоритъм за определяне на най-вероятната дума (или думи), които произвеждат дадената последователност от фонеми.

За щастие, като програмист на Python, не е нужно да се тревожите за нищо от това. Редица услуги за разпознаване на реч са достъпни за използване онлайн чрез API и много от тези услуги предлагат Python SDK .

Избор на библиотеки за разпознаване на реч в Python

В PyPI съществуват няколко библиотеки за разпознаване на реч. Някои от тях включват:

* apiai
* assemblyai
* google-cloud-speech
* pocketsphinx
* SpeechRecognition
* watson-developer-cloud
* wit

Някои от тези библиотеки като wit и apiai – предлагат вградени функции, като обработка на естествен език за идентифициране на намерението на говорещия, които надхвърлят основното разпознаване на реч. Други, като google-cloud-speech, се фокусират единствено върху преобразуването на реч в текст. Има една библиотека, който се откроява по отношение на лекотата на използване която е SpeechRecognition. Разпознаването на реч изисква въвеждане на аудио, а SpeechRecognition прави извличането на този вход наистина лесно. Вместо да се налага да създавате скриптове за достъп до микрофони и обработка на аудио файлове от нулата, SpeechRecognition ще ви накара да стартирате само за няколко минути. Библиотеката SpeechRecognition действа като обвивка за няколко популярни API за реч и по този начин е изключително гъвкава. Един от тях – Google Web Speech API – поддържа API ключ по подразбиране, който е твърдо кодиран в библиотеката SpeechRecognition. Това означава, че можете да станете без да се налага да се регистрирате за услуга. Гъвкавостта и лекотата на използване на пакета SpeechRecognition го правят отличен избор за всеки проект на Python. Въпреки това поддръжката за всяка функция на всеки API, който обвива. И така, сега, след като сте убедени, че трябва да изпробвате SpeechRecognition, следващата стъпка е да го инсталирате във вашата среда.

**Разпознаване на реч**

В днешната ера на бързо напредващи технологии, разпознаването на реч се откроява като нещо обичайно във всеки аспект от живота ни. Разпознаването на реч е лингвистика в компютрите, която позволява на системите да разпознават и превеждат човешка реч в текст. Включва областите на електротехниката, компютърните науки и лингвистиката. Нивото на автоматизация и логика обаче, което виждаме в днешните електронни устройства не би било възможно без концепцията за изкуствен интелект. Използват се гласови асистенти техники за разпознаване на глас, синтез на реч и обработка на естествен език (NLP). Картографирайте гласовите входове към командните изходи. Тези техники са подполе на изкуствените интелект, който позволява на компютърните системи да имат интелект, близък до човешкия (в този случай, по отношение на разбирането на нашата реч). Разпознаването на реч включва няколко стъпки, които включват анализ на звука вълнова форма, разделяне на изказвания чрез мълчание, разпознаване на думите във всяко изказване и произвеждайки крайния резултат под формата на текст (оттук и терминът реч към текст (STT)). За всяко изказване характеристиките се извличат като вектор на характеристиките. Аудио моделът се нарича Hidden Моделът на Марков (HMM) описва всеки последователен процес като речта. В речта разпознаване, акустичният модел, езиковият модел и фонетичният речник са трите използвани неща. В днешния свят разпознаването на реч и гласовите асистенти са навсякъде. от персонализирани гласови записи в кол център, които водят обаждащите се през опциите до бизнеса използвайки имейли и документи, генерирани от говор, тази концепция нарасна много в миналото десетилетие. Гласовите асистенти могат да предоставят голямо разнообразие от услуги и по-специално тези от Amazon Alexa и Google Assistant растат с всеки изминал ден. Те включват предоставяне на информация извлечени от интернет, възпроизвеждане на музика или видеоклипове и заместване на хора в областта на обслужване на клиенти и разговорна търговия. Те също са внедрени в интелигентен гласов асистент в автомобилният сектор за управление на функции и настройки от вътрешността на автомобила чрез гласови команди. Технологичните компании работят за създаването на все по-сложни технология, която ще автоматизира повечето процеси, които извършваме за един денЗа повечето от нас върховният лукс би бил асистент, който винаги ни изслушва вашето обаждане, предвижда всяка ваша нужда и предприема действия, когато е необходимо. Този лукс е сега достъпно благодарение на асистентите с изкуствен интелект, известни още като гласови асистенти. Те влизат донякъде малки пакети и може да извършва различни действия, след като чуе дума за събуждане или команда. Те могат да включват светлини, да отговарят на въпроси, да пускат музика, да правят онлайн поръчки, и т.н. Те също така откриха нарастваща съвместимост с IoT (Интернет на нещата), свързан устройства. Гласовите асистенти не трябва да се бъркат с виртуалните асистенти, които са хора които работят дистанционно и следователно могат да се справят с всякакви задачи. По-скоро гласови асистенти са базирани на технологии. Тъй като гласовите асистенти стават по-здрави, тяхната полезност както в личните и бизнес области също ще растат. Проектът, озаглавен „гласов асистент“, е самостоятелно приложение, което предоставя уникален потребителски интерфейс за изпълнение на ежедневни задачи в системата. Този софтуер може да помогне на потребителите да постигнат задачи, които обикновено биха направили сами. Причината за това е, че изговарянето на командата ще бъде по-лесно и по-бързо от въвеждане или щракване. То предлага потребителски интерфейс на естествен език за клиентите. Приложението може да приеме въвеждане в Интелигентен гласов асистент. Това може да се вземе от микрофона, който повечето устройства вече са вградени в тях. Получава се съответната команда и изпълнен. След като командата приключи изпълнението, се осигурява подходяща обратна връзка продължаване на процеса на разпознаване. Приложението е комбинация от Python и Електронни среди.

Инсталиране на SpeechRecognition

SpeechRecognition е съвместим с Python 2.6, 2.7 и 3.3+, но изисква някои допълнителни стъпки за инсталиране за Python 2. В нашият случай ще го инсталираме на Python 3.11.1

Можем да инсталираме SpeechRecognition от терминал с pip:

pip install SpeechRecognition

Преди да продължите, ще трябва да изтеглите аудио файл. Този, който използвах, за да започна, „harvard.wav.  Уверете се, че сте го запазили в същата директория, в която се изпълнява вашата Python проект

SpeechRecognition прави работата с аудио файлове лесна благодарение на своя удобен AudioFileклас. Този клас може да се инициализира с пътя към аудио файл и предоставя интерфейс на контекстния мениджър за четене и работа със съдържанието на файла.

В момента SpeechRecognition поддържа следните файлови формати:

* WAV: трябва да бъде във формат PCM/LPCM
* AIFF
* AIFF-C
* FLAC: трябва да е оригинален FLAC формат; OGG-FLAC не се поддържа

Цел на проекта

Този проект има за цел да предостави услугите на гласово управление. Приложението може да манипулира системата изпълнява основни задачи чрез разпознаване на реч. В днешния свят на нарастващо търсене за лични асистенти този проект предлага решение, което опростява функционалността на днешни асистенти. Стабилният растеж на тази технология в различни области си доказа своята компетентност в сравнение с автоматизацията, базирана на текст и изображения. Този проект е самостоятелно приложение, което може да се използва от всеки, който не го използва познават работата на системата. Може да изпълнява задачи като отваряне на приложение (като настройки или калкулатор), отваряне на часовника или аларми, контролиране на силата на звука или яркостта и т.н. Тези операции могат да се извършват от произнасяне на командите на глас. Може да анализира и синтезира реч и тук, поддържаният език ще бъде български. Гласовият асистент позволява на потребителите да имат интерфейс със свободни ръце към своите устройства. За потребители, които не знаят как да работят със система, това може да се окаже много полезно. Може да показва уебсайтове, отваряйте други. То ще чуе изречените думи, разпознават думите, преобразуват изречената фраза в текст и след това обработват текущата задача. Машинно изпълнимите команди ще взаимодействат с други приложения в системата и предоставя подходяща обратна връзка на потребителя, който ще бъде уведомен, че задачата е изпълнена успешно. Този проект може да бъде разширен, за да бъде приложен върху различни приложения. Може да се използва на персонални компютри, работещи с всяка операционна система.

Характеристика на проекта

Приложението е базирано на технологията за обработка на естествен говор. Той ще осигури интерфейс със свободни ръце между потребителите и системата. Разпознаването на реч включва превод на изговорени думи или фрази в машинно четим формат чрез анализиране на говорния сигнал. Акустичните етикети са комбинирани с езикови модели за създаване на текстова версия на изговорените команди. Обработката на естествен език се използва за наблюдение и контрол на взаимодействието между хората и компютрите. Той анализира данните от естествения език и помага за трансформирането за въвеждане на реч в текст или обратно. Този проект ще позволи на потребителите да имат достъп и да използват функциите на гласов асистент, като Alexa и Siri. Той ще слуша изговорените команди и ще идентифицира задачите ще бъде изпълнено. Той ще изпълнява операции в системата като отваряне на приложение, търсене онлайн, регулиране на звука, настройване на аларми или отговаряне на запитвания. Дори човек, който не знае как да работи със системата, може лесно и ефективно да получи достъп с помощта на такова приложение. В този проект автоматизацията се извършва с помощта на Python. Този език за програмиране позволява взаимодействие със системата чрез клавишни комбинации, използвайки „pyautogui“ пакет. Чрез програмиране на Python ние също имаме достъп до сървъра на Google, запис, аудио от системния микрофон. Има два различни Pythonфайлове, включени в този проект, един за аудио разпознаване и един за изпълняваща системи задачи. Файловете на Python се изпълняват като дъщерни процеси (процеси на рендеринг) заедно с основния електронен процес, Тези моменти са, когато потребителят взаимодейства със системата или таймерът се изключва.

ГЛАСОВО РАЗПОЗНАВАНЕ

Разпознаването на реч е способността на машината да разпознава думи, изречени в някой език и превежда думите в машинно четим формат. Асистентът е софтуер, който изпълнява задачи на потребителя. Тези асистенти могат да работят с текст, реч или изображения. Гласовите асистенти извършват операции на устройства, базирани наустна команда.

Най-широко използваните гласови асистенти са Siri на Apple, GoogleAssistant, Alexa на Amazon и Cortana на Microsoft. До голяма степен инсталирани са в смартфони, настолни компютри и самостоятелни устройства, тези софтуерни агенти обикновено се интегрират с операционната система и предоставят средство за подобрена достъпност при изпълнение на всяка задача .В началото на 60-те години на миналия век Уилям К. Дерш от IBM създава „Кутия за обувки“, първата в историята реч машина за разпознаване, която можеше да разбере общо шестнадесет думи и да изпълнява математически изчисления, базирани на изречената командна фраза. Предшественикът на днешните системи с гласово управление, имаше хардуер, който се състоеше от микрофон, вкойто потребителят ще произнесе командата. Машината не беше като последните иновации и изискваше всяка дума да бъде ясно, ясно и бавно, с паузи между тях. Въпреки това, следващите 57 години донесе много подобрения в технологиите с навлизането на интернет и облачните изчисления, което доведе до разрастването на виртуалните асистенти, каквито ги познаваме днес. IBMShoebox започна революция със своята технология.

Фиг.2 Първата в историята машина за разпознаване на реч

