

Бобрешов Алексей 

UNIVERSITY
OF ARTIFICIAL
INTELLIGENCE

Голосовое управление системой “Умный дом”

Дипломный руководитель
Лукоянов Михаил

2020-2021г.

Группа „Pro“

Идея использования нейросети.



Система «Умный дом» достаточно популярна в разных странах.

Современные системы работают посредством прямого выбора команды на пульте, на панели/ консоли управления, смартфоне и т. д., некоторые варианты имеют систему голосовых команд, которые вводятся в систему (при установке настройке), а после активации «системы» ждут совпадения после произнесения ключевого слова, тем самым инициализируя алгоритм записи и сравнения голосовой команды ранее сохраненной в системе.

Задача: создать нейросеть которая будет прослушивать звуки в доме/квартире и вычлняя из речи команды, исполнять их.

Так-же систему можно снабдить возможностью выявлять аномалии в голосе для дальнейшего использования как в целях защиты так и для наблюдения за здоровьем.

Цель



- Научить нейросеть определять голосовые команды во время обычного разговора в помещении
- Научить определять тревожность в голосе и выявлять аномалии от обычного диалога
- Определять детские голоса и эмоциональное состояние обитателей по нескольким категориям
- Понять возможности развития данного направления искусственного интеллекта в том числе в смежном применении

Постановка задачи

- Сбор голосовых команд на управление умными устройствами
- Структурирование данных (команды, повседневная речь)
- Написание модели нейросети и тренировка
- Проверка качества обучения

Нейросеть «Умный дом», исключит зависимость от консоли управления, не будет требовать высоких мощностей и модернизации оборудования, А так-же прочих механических изменений компонентов продукта от конечного пользователя. При увеличении количества «умных приборов», даст возможность расширять количество команд управления при появлении новых «умных устройств» в помещении, а так-же будет иметь возможность в несколько вариантов продукта:

заботливый помощник будет подсказывать\напоминать о наступлении событий, управление устройствами умного дома при их подключённости к системе «Умного дома», не используя дополнительные приспособления для подачи команд и при этом с возможностью разделения полномочий по возрасту (дети / взрослые), безопасность жилья с определением присутствия посторонних, выявление паттернов поведения каждого индивидуума дома, определение эмоционального состояния для ободряющей музыки или напитка который поднимет настроение.

Использование настроек различной техники, при возможности подключения к управляющей системе данной техники. Что в свою очередь будет способствовать развитию самих умных устройств как во взаимодействии с ИИ (искусственным интеллектом) так и самостоятельно. Например можно начать производство небольших управляемых узлов (управляемые выключатели) и развивать к более умным устройствам.

Идея проста: заходя в помещение можно включить лампу механически (выключателем), можно установить датчик движения который не требует свободной руки, но при отсутствии движения отключит освещение, а ИИ не даст это сделать, вернее пока не получит команду Выключи свет в прихожей или уменьши яркость ещё и ещё.. не оставит без освещения, а при наличии дневного (солнечного) света, даже не будет включать.

Вы просто подаёте голосовую команду на выполнение действия, система определяет её и исполняет.

Развитие системы охраны помещения. Умная сигнализационная система охраны при выявлении опасных ситуаций — будет реагировать подавая сигнал бедствия в соответствующие службы.

Распознавая присутствие хозяев помещения и гостей, умная система доступа может контролировать и сообщать о не званых посетителях, как при обычном освещении так и при инфракрасном, включая определение звуков особенно на предприятиях под охраной, подавая сигнал при несанкционированном доступе на пульт дежурного и(или) правоохранительных органов, при этом различая человека и мелкое животное.

Возможно определение тревожности в голосе или поведении — позволит решать более важные задачи, такие как определение болезненного состояния, предлагая варианты для восстановления, а если это , так-же при определении нежелательного контента «Умный дом» сможет ограничивать его если это происходит в присутствии детей. так-же при детекции огня в помещении или совершенно бытовых задачах когда нейросеть отследит .

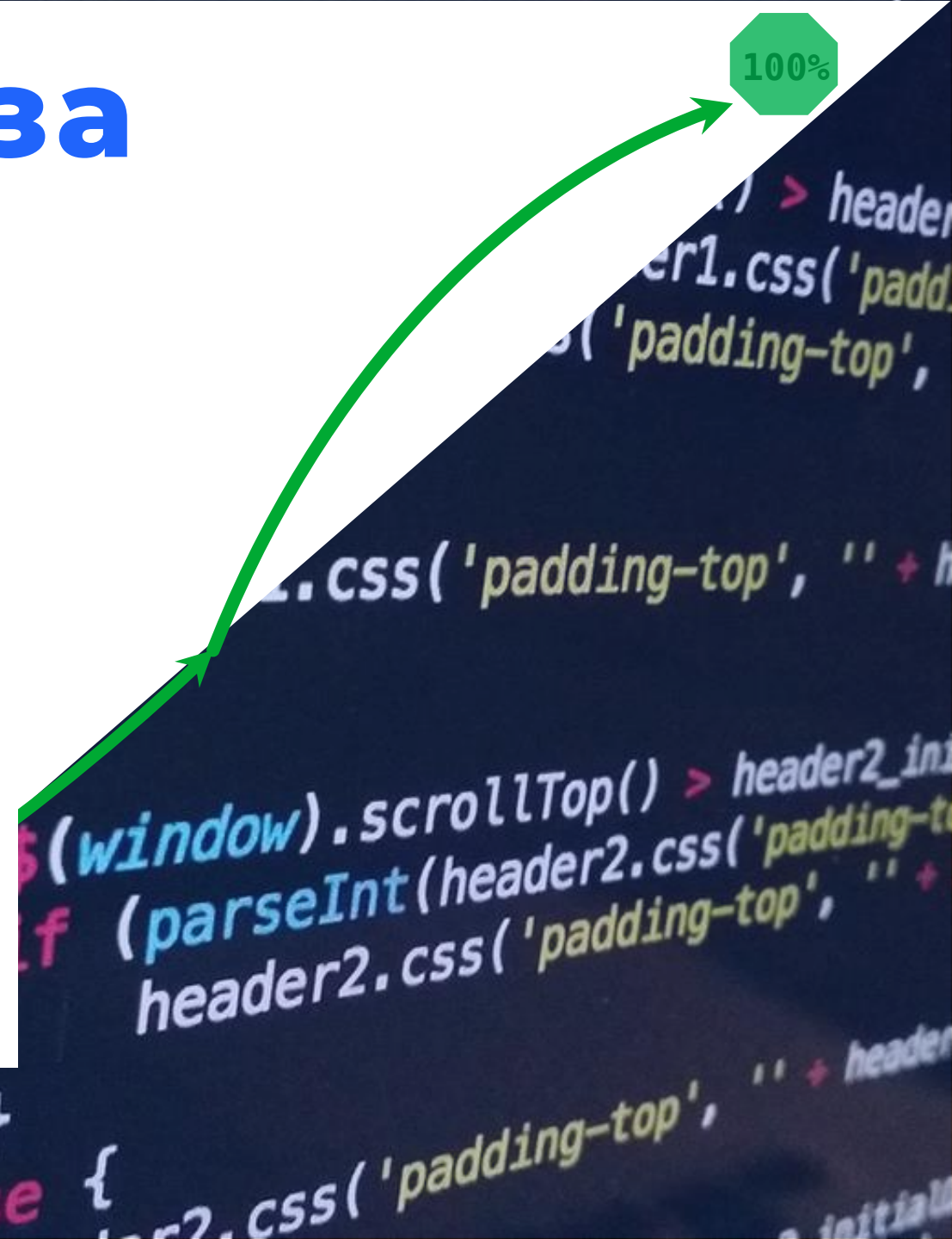
Реализация

- Сбор данных для обучения модели
(в моём случае это набор голосовых команд для извлечения признаков и голосовое чтение текста для создания признаков фона)
- Разметка полученных данных, обработка и извлечение признаков ...
- Составление оптимальной модели нейросети, которая не нагружая систему будет определять команды из общей речи
- Проверка работоспособности
получение первых результатов точности

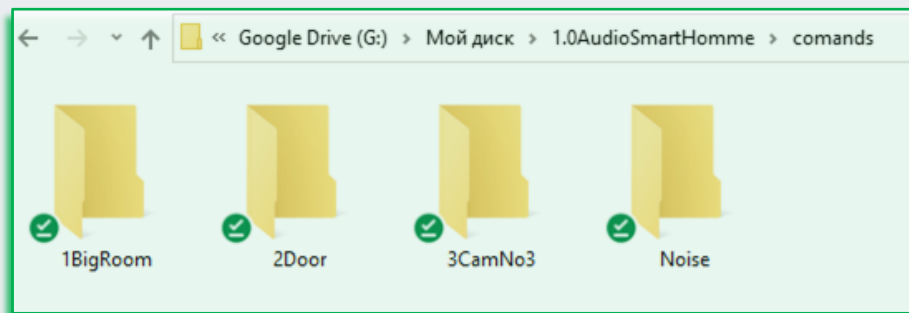


Обучающая база

- ✓ Объем базы
начал с 620 звуковых файлов
- ✓ Источник сбора данных
человек (голос), фоновые звуки (телевизор новости и чтение книги)
- ✓ Инструменты для сбора данных
Диктофон/смартфон лежащий на равном удалении от источников звука
- ✓ Возникшие трудности/недочёты
Желательно иметь различные голоса для тренировки (возрастное различие, по гендерное различие, региональные различия)
- ✓ Тренировка в среде GoogleColab внесла ограничения применение слоёв DL (глубокого обучения) по причине небольшого выделения ресурсов



Сбор и структурирование данных



bigROOM (10).wav	00:00:04	751 КБ
bigROOM (29).wav	00:00:03	691 КБ
bigROOM (31).wav	00:00:03	691 КБ
bigROOM (33).wav	00:00:03	691 КБ
bigROOM (35).wav	00:00:03	721 КБ
bigROOM (37).wav	00:00:03	721 КБ
bigROOM (39).wav	00:00:03	721 КБ
bigROOM (41).wav	00:00:03	571 КБ
bigROOM (43).wav	00:00:03	571 КБ
bigROOM (45).wav	00:00:03	571 КБ
bigROOM (46).wav	00:00:02	421 КБ
bigROOM (47).wav	00:00:02	421 КБ
bigROOM (48).wav	00:00:02	421 КБ
bigROOM (49).wav	00:00:05	1 021 КБ
bigROOM (50).wav	00:00:05	1 021 КБ
bigROOM (51).wav	00:00:05	1 021 КБ
bigROOM (52).wav	00:00:04	871 КБ
bigROOM (53).wav	00:00:04	871 КБ
bigROOM (54).wav	00:00:04	871 КБ
bigROOM (55).wav	00:00:04	901 КБ
bigROOM (56).wav	00:00:04	901 КБ
bigROOM (57).wav	00:00:04	901 КБ
bigROOM (58).wav	00:00:04	781 КБ
bigROOM (59).wav	00:00:04	781 КБ

Создаём набор голосовых команд, путём произношения команд управления на записывающее устройство.

Переносим данные в облачный сервис для дальнейшей работы с ним.

Создаём папки с названиями команд и сортируем соответствующим образом.



Выбор модели нейросети

Первично была использована сеть прямого распространения, она дала очень хороший результат, но при увеличении данных и похожести голосов все результаты изменились на противоположные.

Следующей моделью была выбрана CNN (свёрточная сеть Conv1D). Она показала хороший результат на различных объёмах данных и различных параметрах обучения, значительно менялись результаты при изменении активационных функций.

В связи с вышеперечисленным принято решение оставить несколько вариантов обработки данных для одних и тех-же баз.

Нейронка с несколькими входами будет иметь возможность вычислять разные признаки в одном наборе данных, и при добавлении новых в том числе.

Данная сеть очень хорошо себя показала в обучении на разных наборах данных, именно в скорости и точности. Это важный момент, т. к. в дальнейшем планируется сильно расширять возможности, по мере подключения новых устройств к управлению, возможно будут новые люди, а следовательно база данных для обучения будет существенно расти.

```
Epoch 45/100
4155/4155 [=====] - 1s 175us/step - loss: 0.01220000 - accuracy: 0.9966 - val_loss: 0.4374 - val_accuracy: 0.9234
.....
Epoch 83/100
4155/4155 [=====] - 1s 178us/step - loss: 5.9840e-05 - accuracy: 1.0000 - val_loss: 0.3459 - val_accuracy: 0.9406
.....
Epoch 100/100
4155/4155 [=====] - 1s 176us/step - loss: 2.5689e-05 - accuracy: 1.0000 - val_loss: 0.6442 - val_accuracy: 0.9109
```

На этих данных видим обучение сети, количество Эпох, время обучения (расчета) за одну эпоху и каждого шага в эпохе, коэффициент ошибки сети за эпоху и коэффициент точности на обучающихся данных, в крайних двух видим данные по проверочной выборке:

Коэффициент ошибки - val-loss

Коэффициент точности - val-accuracy

Обучение на
100 эпохах

Эволюция нейросети SmartHome

На различных этапах развития технологий и техники можно будет увеличивать возможности



Связь и контроль устройств



Умение предугадывать состояния



Интеллектуальное управление

ДОМ КОТОРЫЙ ТЕБЯ ПОНИМАЕТ

На данный момент из открытых источников, библиотек, для определения интеллектуального распознавания голоса не много. Возможностей по идентификации и определению состояния человека в различные моменты времени ещё нет, но то что это можно делать — факт. Развитие искусственного интеллекта в этом направлении даст широкие возможности для различных областей науки и техники, улучшая жизнь человека в целом и оберегая его в различных жизненных ситуациях.

Распознавание команд из обычной речи — это начало которое позволит сделать жизнь комфортнее.

А предприятию начавшему развивать данное направление, даст возможность создать целую эпоху умных устройств.