

# РОЗРОБКА ERP-ЗАСТОСУНКУ ДЛЯ ГОЛОСОВОГО УПРАВЛІННЯ ТИПОВИМИ ОПЕРАЦІЯМИ

<sup>1</sup> Вінницький національний технічний університет;

## Анотація

Авторами розроблено програму, що має змогу розпізнавати ключові слова у потоці мовлення. Застосування написано на мові програмування Python з використанням бібліотеки Sphinx. Створене застосування сприймає граматику, словник, та реагує на набір голосових команд, що може бути вказаний користувачем. Окремо реалізований також CLI інтерфейс для створеного застосування.

**Ключові слова:** голосового управління, розпізнавання мови.

## Abstract

The authors have developed a program that can recognize keywords in the speech stream. The application is written in the Python programming language using the Sphinx library. The created application perceives grammar, vocabulary, and responds to a set of voice commands that can be specified by the user. The CLI interface for the created application is also separately implemented.

**Keywords:** server voice control, speech recognition.

## Вступ

Нині все частіше використовується системи розпізнавання голосових команд в потоці мовлення. Дана технологія практично незамінна для створення додаткового каналу управління в різних автоматизованих системах. Проте, в даний час для української мови безкоштовних робочих рішень немає. Хоча існує безліч алгоритмів для розпізнавання голосових команд, більшість з них засновані на прихованих марковських моделях або нейромережах. Приховані марковські моделі базуються на імовірнісному співвідношенні порядку фонем у слові. Вони добре працюють при невеликому розмірі словника, але при розмірі словника більше певної межі цей метод практично перестає працювати. Таким чином, розробка методу, здатного розпізнавати ключові слова в потоці мовлення з великим розміром словника є актуальним завданням. Для певного набору прикладних задач можна використовувати CMUSphinx – відкритий пакет для роботи з системами мовлення. Для роботи з ним треба вибрати акустично модель. На жаль, в вільному доступі моделі для української мови немає, проте є інші, які мають схожі властивості.

Метою є дослідження методів розпізнавання голосових команд в потоці мовлення і розробці програмного продукту для розпізнавання голосових команд українською мовою.

## Результати дослідження

Опишемо процес створення та тестування додатку для розпізнавання голосових команд української мови, заснованого на бібліотеці PocketSphinx. PocketSphinx можна використовувати як Python пакет або як системну програму CLI. Наведемо приклад Python використання PocketSphinx пакету для розпізнавання голосових команд (Рис. 1). У якості параметрів маємо: hmm – шлях до акустичної моделі, keyphrase – ключове слово, dict – фонетичний словник.

```
pocketsphinx_dir=os.path.dirname(pocketsphinx. file )
model_dir = os.path.join(pocketsphinx_dir, 'model') config = pocketsphinx.Decoder.default confi
onfig.set_string('-hmm',os.path.join(model_dir,'rus')) config.set_string('-keyphrase','привіт')
config.set_string('-dict',os.path.join(model_dir,dict_name)) config.set_float('-kws_threshold',self.threshold)
p = pyaudio.PyAudio()
stream = p.open(format=pyaudio.paInt16, channels=1, rate=16000, input=True, frames_per_buffer=20480)
stream.start_stream() buf = stream.read(1024) if buf:
decoder.process_raw(buf)
else:
break;
if decoder.hyp() is not None: print("Hotword Detected") decoder.end_utt() start_speech_recognitio
ecoder.start_utt()
```

Рис. 1. Реалізація системи розпізнавання голосових команд на базі PocketSphinx

Фонетичний словник зставляє слово словника до відповідної послідовності фонем. Це може виглядати так: привіт -> P R I V I T T. Так ми зробили скрипт, що здатний розпізнавати слово «привіт». Так ми отримали готове рішення для розпізнавання голосових команд у потоці мовлення для української мови. Проте, цей підхід має декілька недоліків: - ми можемо використовувати лише Python у якості мови програмування; - Pip версія PocketSphinx не має останні оновлень продукту. Отже, також зроблено власне більш універсальне рішення. Для цього, спочатку треба скомпілювати SphinxBase та PocketSphinx під відповідну платформу. Так, отримуємо вихідний код з офіційного репозиторію PocketSphinx за допомогою git та відповідної команди клонування (Рис. 2).

```
git clone https://github.com/cmusphinx/pocketsphinx.git
```

Рис. 2. Отримання коду PocketSphinx за допомогою git

Далі запустимо процес з установки з офіційних bash скриптів (Рис. 3). Для установки нам знадобиться встановлений C компілятор та make.

```
./configure
make clean all
make check
make
```

Рис. 3. Команди компіляції PocketSphinx

Тепер можна використовувати засоби CLI PocketSphinx такі як pocketsphinx\_continuous. Запустимо у CLI режимі команду для розпізнавання голосових команд у потоці мовлення. Для цього створимо словник та додамо туди слово “відкрити”, що має фонетичну форму v i d k r u t y. Далі, вже можемо виконати команду для початку розпізнавання (Рис. 4). Де hmm - шлях до акустичної моделі, dict - шлях до словника, а keyphrase - задане ключове слово.

```
pocketsphinx_continuous -hmm zero_ru.cd cont_4000
inmic yes -keyphrase відкрити -dict wordlist.voc
```

Рис. 4. Команда запуску pocketsphinx\_continuous для розпізнавання голосових команд у потоці мовлення

Після запуску команди початку аналізу можемо бачити логи програми, та відповідні розпізнані слова (Рис. 5).

```
INFO: continuous.c(275): Ready...
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(120): Update from < 9.33 -0.01 -0.13 -0.06 -0.25 -0.20 -0.18 -0.13 -0.18 -0.14 -0.15 -0.15 -0.09 >
INFO: cmn_live.c(138): Update to < 9.97 -0.10 -0.17 -0.05 -0.31 -0.21 -0.16 -0.15 -0.22 -0.14 -0.13 -0.15 -0.09 >
INFO: kws_search.c(656): kws 0.44 CPU 0.181 xRT
INFO: kws_search.c(658): kws 3.73 wall 1.529 xRT
INFO: continuous.c(275): Ready...
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 9.97 -0.10 -0.17 -0.05 -0.31 -0.21 -0.16 -0.15 -0.22 -0.14 -0.13 -0.15 -0.09 >
INFO: cmn_live.c(105): Update to < 9.37 -0.09 -0.13 -0.08 -0.28 -0.23 -0.18 -0.14 -0.19 -0.12 -0.12 -0.14 -0.09 >
INFO: cmn_live.c(88): Update from < 9.37 -0.09 -0.13 -0.08 -0.28 -0.23 -0.18 -0.14 -0.19 -0.12 -0.12 -0.14 -0.09 >
INFO: cmn_live.c(105): Update to < 9.36 -0.16 -0.05 -0.10 -0.26 -0.26 -0.18 -0.13 -0.20 -0.12 -0.11 -0.14 -0.09 >
INFO: cmn_live.c(120): Update from < 9.36 -0.16 -0.05 -0.10 -0.26 -0.26 -0.18 -0.13 -0.20 -0.12 -0.11 -0.14 -0.09 >
INFO: cmn_live.c(138): Update to < 9.62 -0.14 -0.05 -0.08 -0.29 -0.26 -0.18 -0.14 -0.20 -0.12 -0.11 -0.14 -0.09 >
INFO: kws_search.c(656): kws 1.12 CPU 0.185 xRT
INFO: kws_search.c(658): kws 7.44 wall 1.229 xRT
відкрити
```

Рис. 5. Логи pocketsphinx\_continuous

## Висновки

Авторами розроблено програму, що має змогу розпізнавати ключові слова у потоці мовлення. Застосування написано на мові програмування Python з використанням бібліотеки Sphinx. Створене застосування сприймає граматику, словник, та реагує на набір голосових команд, що може бути вказаний користувачем. Окремо реалізований також CLI інтерфейс для створеного застосування.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Тампель И.Б., Карпов А.А. Автоматическое распознавание речи. Учебное пособие. — СПб.: Университет ИТМО, 2016. — 138 с.
2. Пикон Дж. Методы моделирования сигнала в распознавании речи. Перев. Попова Р., Кемерово, 2000. — 79 с.
3. Кипяткова И.С., Ронжин А.Л., Карпов А.А. Автоматическая обработка разговорной русской речи. Санкт-Петербургский институт информатики и автоматизации Российской Академии Наук, 2013, — 316 с.

*Дихніч Людмила Дмитрівна* — студентка групи 2АКІТ-20м, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: malyishka4@gmail.com

*Майданевич Ірина Олександрівна* — студентка групи 2АКІТ-20м, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: maidanevych.i.99@gmail.com

Науковий керівник: *Ковтун В'ячеслав Васильович* — д-р техн. наук, доцент, професор кафедри комп'ютерних систем управління, Вінницький національний технічний університет, м. Вінниця

*Dykhnich Lyudmyla D.* – student of group 2AKIT-20m, faculty of computer systems and automation, Vinnytsia National Technical University, Vinnytsia, e-mail: malyishka4@gmail.com

*Maidanevych Iryna O.* – student of group 2AKIT-20m, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: maidanevych.i.99@gmail.com

Supervisor: *Kovtun Viacheslav V.* - Dr. Tech. Sciences, Associate Professor, Professor of Computer Control Systems Department, Vinnytsia National Technical University, Vinnytsia