Speech Recognition

- With Information Retrieval & Command Prompt -

Project Report

ED5-1-E17

Aalborg University
Electronics and IT

Copyright © Aalborg University 2015

Here you can write something about which tools and software you have used for typesetting the document, running simulations and creating figures. If you do not know what to write, either leave this page blank or have a look at the colophon in some of your books.

Electronics and IT
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Speech Recognition

Theme:
Scientific Theme

Project Period:
Fall Semester 2017

Project Group:
ED5-1-E17

Participant(s):
Stefan Bîrs
Tiberiu-Ioan Szatmari
Kamran Thomas Alimagham

Supervisor(s):
Daniel Ortiz-Arroyo
D. M. Akbar Hussain

Copies: 1

Page Numbers: 19

Date of Completion:
October 23, 2017

Abstract:

This report goes into detail about the implementation of distributed speech recognition systems, information retrieval and machine learning. The main objective of the speech recognition system is to provide a handsfree human-computer interface experience, where the system is also able to handle noise.

The content of this report is freely available, but publication (with reference) may only be pursued due to agreement with the author.

Contents

Preface **ix**

1 Introduction **1**
1.1 Examples . . . 1
1.2 How Does Sections, Subsections, and Subsections Look? . . . 1
1.2.1 This is a Subsection . . . 1

2 Problem Description **3**
2.1 Problem Description . . . 3
2.2 Problem Delimitation . . . 3

3 Speech Processing **5**
3.1 The basic units that make up the sounds of words . . . 5

4 Machine Learning for Speech **7**
4.1 TensorFlow . . . 7
4.1.1 Tensors . . . 7
4.1.2 TensorFlow Core . . . 8
4.1.3 tf.estimator . . . 8
4.2 Deep neural network for speech . . . 8
4.2.1 Deep Feed Forward (DFF) . . . 8
4.2.2 Recurrent Neural Nerwork (RNN) . . . 8

5 Distributed speech recognizer **9**
5.1 The DSR system . . . 9
5.2 Acoutic models . . . 9
5.3 Language model . . . 9

6 Knowledge-based IR system **11**
6.1 The document retrieval system . . . 11
6.2 The question answering system . . . 11

7 Discussion **13**

8 Conclusion **15**

Bibliography **17**

A Appendix A name **19**

Todo list

Is it possible to add a subsubparagraph? 2
I think that a summary of this exciting chapter should be added. 2
Found on page 396 5
reference this part with *https://www.allerin.com/blog/six−types−of−neural−networks* 8
link this part with link above 8

Preface

The project entitled Speech Recognition with Focus on information retrieval and command prompt was made by three students from the Electronics and Computer Engineering programme at Aalborg University Esbjerg, for the P5 project during the fifth semester.

Aalborg University, October 23, 2017

Stefan Bîrs
<sbars15@student.aau.dk>

Tiberiu-Ioan Szatmari
<tszatm15@student.aau.dk>

Kamran Thomas Alimagham
<kalima15@student.aau.dk>

Chapter 1

Introduction

Here is the introduction. The next chapter is chapter 1.
a new paragraph

1.1 Examples

You can also have examples in your document such as in example 1.1.

Example 1.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \,. \tag{1.1}$$

You can adjust the colour and the line width in the `macros.tex` file [1].

1.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

1.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this [2].

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph title [3].

Is it possible to add a subsubparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 2

Problem Description

2.1 Problem Description

List to wright about here:

- Main Objective/Goal
- What will be required
- How are we going solve the Objective?

2.2 Problem Delimitation

To ensure that the project meets the desired scope, specific success criteria are formed which will be evaluated during the final stages of the report. The following list of requirements are the desired success criteria, where the focus will ve on reaching the given goals.

- Detect human speech
- Design a deep neural network
- Compare different regulation technics (drop out, bach normalization...)
- Design a filter
- Create a network infrastructure for Client/Server Model (Distribution)
- Implement a speech recognition system with information retrival

Brief summary on how are we going to obtain all of these bullet points.

Chapter 3

Speech Processing

3.1 The basic units that make up the sounds of words

Found on page 396

Chapter 4

Machine Learning for Speech

4.1 TensorFlow

TensorFlow provides multiple Application Programming Interfaces (APIs) for machine learning. The lowest level API "TensorFlow Core" provides complete programming control [4]. For those who require fine levels of control over their models, TensorFlow Core is a well-suited tool for the job. There are higher level APIs that are built on top of TensorFlow Core. These higher level APIs are typically easier to learn and use than Tensorflow Core [4]. In addition, the higher level APIs such as "tf.estimator" helps with managing data sets, estimators, training and inferences (testing your trained network), as well as, making repetitive tasks easier and more consistent [4].

The subsection below will start with TensorFlow Core. In order to gain an understanding of the basic principles that TensorFlow has to offer, a model shall be made. In the subsection after that, the same model will be implemented with tf.estimator. Knowing TensorFlow Core principles will give us a great mental model of how things are working internally when we use the more compact higher level API.

4.1.1 Tensors

The central unit of data in TensorFlow is the tensor. A tensor consists of a set of primitive values shaped into an array of any number of dimensions. A tensor's rank is its number of dimensions. Here are some examples of tensors:

```
3 # a rank 0 tensor; this is a scalar with shape []
[1., 2., 3.] # a rank 1 tensor; this is a vector with shape [3]
[[1., 2., 3.], [4., 5., 6.]] # a rank 2 tensor; a matrix with shape [2, 3]
[[[1., 2., 3.]], [[7., 8., 9.]]] # a rank 3 tensor with shape [2, 1, 3]
```

4.1.2 TensorFlow Core

4.1.3 tf.estimator

4.2 Deep neural network for speech

4.2.1 Deep Feed Forward (DFF)

The simplest of all neural networks, the feedforward neural network, moves information in one direction only. Data moves from the input nodes to the output nodes, passing through hidden nodes (if any). The feedforward neural network has no cycles or loops in its network.

reference this part with *https://www.allerin.com/blog/six-types-of-neural-networks*

4.2.2 Recurrent Neural Nerwork (RNN)

The recurrent neural network, unlike the feedforward neural network, is a neural network that allows for a bi-directional flow of data. The network between the connected units forms a directed cycle. Such a network allows for dynamic temporal behavior to be exhibited. The recurrent neural network is capable of using its internal memory to process arbitrary sequence of inputs. This neural network is a popular choice for tasks such as handwriting and speech recognition.

link this part with link above

Long/Short Term Memory (LSTM)

Chapter 5

Distributed speech recognizer

5.1 The DSR system

5.2 Acoutic models

5.3 Language model

Chapter 6

Knowledge-based IR system

6.1 The document retrieval system

6.2 The question answering system

Chapter 7

Discussion

Chapter 8

Conclusion

In case you have questions, comments, suggestions or have found a bug, please do not hesitate to contact me. You can find my contact details below.

Jesper Kjær Nielsen
jkn@es.aau.dk
http://kom.aau.dk/~jkn
Fredrik Bajers Vej 7
9220 Aalborg Ø

Bibliography

[1] Tobias Oetiker. *The Not So Short A Introduction to LaTeX2e*. http://tobi.oetiker.ch/lshort/lshort.pdf. 2010.

[2] Lars Madsen. *Introduktion til LaTeX*. http://www.imf.au.dk/system/latex/bog/. 2010.

[3] Andrew S Tanenbaum and Albert S Woodhull. *Operating systems: design and implementation*. Vol. 2. Prentice-Hall Englewood Cliffs, NJ, 1987.

[4] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

Appendix A

Appendix A name

Here is the first appendix