

# Learning a Disease Embedding using Generalized Word2Vec Approaches.

Milan van der Meer

Thesis submitted for the degree of  
Master of Science in Engineering:  
Computer Science, specialisation  
Artificial Intelligence

**Thesis supervisor:**

Prof. dr. Roel Wuyts

**Assessors:**

Ir. Kn. Owsmuch

K. Nowsrest

**Mentors:**

Ir. An Assistent

S. Dhooghe

© Copyright KU Leuven

Without written permission of the thesis supervisor and the author it is forbidden to reproduce or adapt in any form or by any means any part of this publication. Requests for obtaining the right to reproduce or utilize parts of this publication should be addressed to the Departement Computerwetenschappen, Celestijnenlaan 200A bus 2402, B-3001 Heverlee, +32-16-327700 or by email [info@cs.kuleuven.be](mailto:info@cs.kuleuven.be).

A written permission of the thesis supervisor is also required to use the methods, products, schematics and programs described in this work for industrial or commercial use, and for submitting this publication in scientific contests.

# Preface

I would like to thank everybody who kept me busy the last year, especially my promotor and my assistants. I would also like to thank the jury for reading the text. My sincere gratitude also goes to my wife and the rest of my family.

*Milan van der Meer*

# Contents

<b>Preface</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>List of Figures and Tables</b>	<b>iv</b>
<b>List of Abbreviations and Symbols</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Lorem Ipsum 4–5 . . . . .	1
1.2 Lorem Ipsum 6–7 . . . . .	1
<b>2 Background</b>	<b>3</b>
2.1 Introduction . . . . .	3
2.2 Background Knowledge . . . . .	3
2.3 Word2Vec . . . . .	7
2.4 DeepWalk . . . . .	9
2.5 Conclusion . . . . .	10
<b>3 The Final Chapter</b>	<b>11</b>
3.1 The First Topic of this Chapter . . . . .	11
3.2 The Second Topic . . . . .	12
3.3 Conclusion . . . . .	13
<b>4 Conclusion</b>	<b>15</b>
<b>A The First Appendix</b>	<b>19</b>
A.1 More Lorem . . . . .	19
A.2 Lorem 51 . . . . .	20
<b>B The Last Appendix</b>	<b>21</b>
B.1 Lorem 20-24 . . . . .	21
B.2 Lorem 25-27 . . . . .	22
<b>Bibliography</b>	<b>23</b>

# Abstract

The **abstract** environment contains a more extensive overview of the work. But it should be limited to one page.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

# List of Figures and Tables

## List of Figures

2.1	A simple presentation of a perceptron. . . . .	4
2.2	A more complex network made by connecting multiple perceptrons. . .	4
2.3	General vocabulary for a multilayer network. . . . .	5
2.4	A small change on the weights, only has a small impact on the output. .	5
2.5	Visual representation of the terminology for a neural network. . . . .	6
2.6	Explanation of n-gram. . . . .	8
2.7	Overview of the DeepWalk algorithm. . . . .	10

## List of Tables

# List of Abbreviations and Symbols

## Abbreviations

LoG	Laplacian-of-Gaussian
MSE	Mean Square error
PSNR	Peak Signal-to-Noise ratio

## Symbols

42	“The Answer to the Ultimate Question of Life, the Universe, and Everything” according to <a href="#">[1]</a>
$c$	Speed of light
$E$	Energy
$m$	Mass
$\pi$	The number pi





# Chapter 1

## Introduction

The first contains a general introduction to the work. The goals are defined and the modus operandi is explained.

### 1.1 Lorem Ipsum 4–5

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

### 1.2 Lorem Ipsum 6–7

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

# Chapter 2

## Background

### 2.1 Introduction

In this chapter

### 2.2 Background Knowledge

We assume a basic knowledge of general computer science concepts as algorithms, statistics, time complexity, linear algebra, basic graph theory, optimization, and heuristics.

#### 2.2.1 Machine Learning

Machine learning is a data driven approach with as goal to build a model which can be used to make predictions or decisions. This task is done by algorithms which are able to learn models based on examples given by the designer. In machine learning there are 3 types of problems, namely supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning is concerned with the learning task where there are examples given with their corresponding label. Unsupervised learning is similar to supervised learning only no labels are given. We won't go into reinforcement learning.

We can also classify to problems according to the desired output of our model. Those main tasks consist of classification, regression, clustering, and dimension reduction.

#### 2.2.2 Time Series Analysis

A time series consists of data points over a certain time period. We refer to this as a sequence of states. Where a state represents a data point and can differ from a single value to more complex representations like pictures.

The domain of time series analysis handles around extracting information or relations from a time series. It can have different goals like forecasting, classification, or exploratory.

### 2.2.3 Neural Networks

A neural network is a machine learning approach based on biological neural networks.

#### Perceptron

The basic component of a neural network is a perceptron. A perceptron take multiple binary inputs and has a single binary output (see figure 2.1). Each input has a corresponding real numbered weight. The output is decided on the following equation:

$$output = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{threshold} \end{cases} \quad (2.1)$$

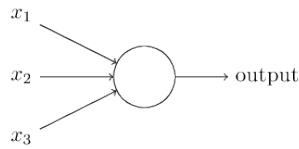


FIGURE 2.1: A simple presentation of a perceptron.

We can build a network by connecting multiple perceptrons (see figure 2.2. By building these networks, more complex decisions can be made.

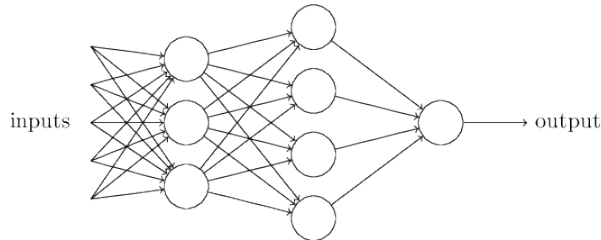


FIGURE 2.2: A more complex network made by connecting multiple perceptrons.

Now we have seen how a general network is constructed, we look at some vocabulary.

In figure 2.3, we see a four-layer network. As mentioned on the figure, we call the first layer the input layer, the last layer the output layer, and the layers in between are called hidden layers. Sometime multiple layer networks are referred to as multilayer perceptrons or MLPs.

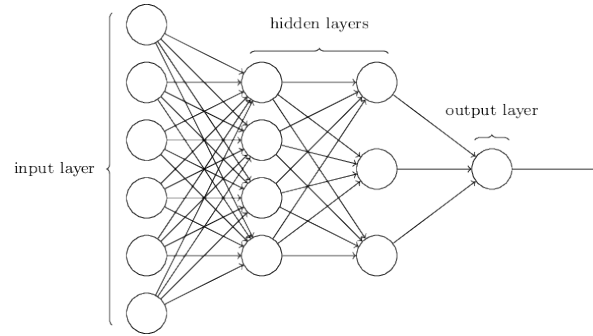


FIGURE 2.3: General vocabulary for a multilayer network.

### Training a network

To train a neural network, we input an example with a known label. The network will have a certain output based on the current weights. When this output is incorrect, it should be possible to adjust the weights with as effect that the network now has as output the correct label. Of course, this change in weights, should only effect the output by a small bit (see figure 2.4). The reason for this is that otherwise all the previous images could now be labeled incorrectly. So, the concept of training a neural network means, adjusting the weights in a way that the behavior of the network doesn't change completely on the previous seen pictures but that the current picture is labeled correctly.

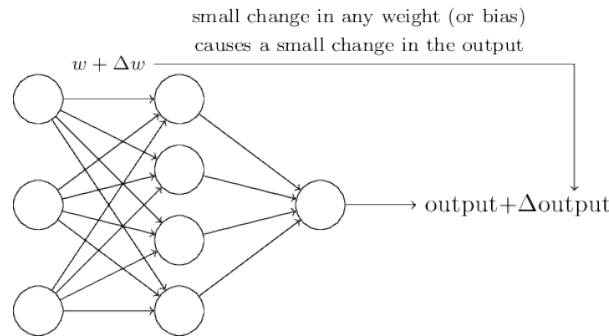


FIGURE 2.4: A small change on the weights, only has a small impact on the output.

To achieve this effect, we change our known perceptrons to sigmoid neurons. A sigmoid neuron has the same basics as a perceptron. It still has inputs but now it also has a bias  $b$ . The inputs still have weights but the weights can now range between 0 and 1. The output is now calculate with  $\sigma(w * x + b)$  where  $\sigma$  is the sigmoid function. This results in the following formula:

$$\frac{1}{1 + \exp(-\sum_j w_j x_j - b)} \quad (2.2)$$

## 2. BACKGROUND

---

The sigmoid function makes it possible to calculate the gradient and makes the output a linear combination of  $\Delta w_j$  and  $\Delta b$  as  $\Delta output$  is approximated by

$$\Delta output \approx \sum_j \frac{\partial output}{\partial w_j} \Delta w_j + \frac{\partial output}{\partial b} \Delta b \quad (2.3)$$

Because of the linearity, it is now possible to choose changes for the weights and biases to achieve a correct output. By adjusting the weights, we will train our network to achieve a higher accuracy.

### 2.2.4 Backpropagation

Backpropagation is an algorithm which is used to train neural networks. It calculates the gradient of a chosen loss function with respect to the individual weights. With the gradient, the weights are updated and the loss function is minimized.

#### Terminology

We use  $w_{jk}^l$  to denote the weight corresponding to the connection between the  $k^{th}$  node in the  $(l-1)^{th}$  layer to the  $j^{th}$  node in the  $l^{th}$  layer. We use  $b_j^l$  for the bias of the  $j^{th}$  node in the  $l^{th}$  layer and  $a_j^l$  for the activation of the  $j^{th}$  node in the  $l^{th}$  layer. See figure 2.5.

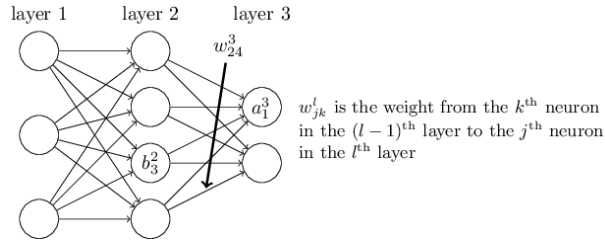


FIGURE 2.5: Visual representation of the terminology for a neural network.

We can now convert these notation to a vector representation. We remove the indexes for the node numbers which results in the following:

$$a^l = \sigma(w^l a^{l-1} + b^l) \quad (2.4)$$

#### Cost function

AS mentioned before, backpropagation has as goal to calculate the partial derivatives of the cost function  $C$  with respect to each weight and bias.

The cost function has to fulfill certain criteria. The first one is that it needs to be possible to write it as a summation over cost functions for individual training examples. Secondly, it needs to be derivable. And lastly, the cost function is a function of the activations of the last layer.

### Fundamental equations

Backpropagation has 4 equations. They allow us to calculate the error for each node and adjust the weights based on the gradient descent.

First we calculate the error of each node which is based on how much the cost function is influenced by each activation and on how much the activation function is influenced by  $z_j^L$ :

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} \sigma'(z_j^L) \text{ with } z_j^L = \sum_k w_{jk}^L a_k^{L-1} + b_j^L \quad (2.5)$$

This can be written as a neat vector equation:

$$\delta^L = (a^L - y) \circ \sigma'(z_j^L) \quad (2.6)$$

The next equation explains why the algorithm is called backpropagation. The equation calculates each layers error vector based on the layer after it, it propagates the error back over the layers:

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \circ \sigma'(z_j^L) \quad (2.7)$$

With those 2 equations we calculate the error in each layer of the neural network. Those errors can be used to calculate the derivatives of the cost function with respect to the weights and the biases:

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l. \quad (2.8)$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l. \quad (2.9)$$

When the derivatives are calculated, we can apply the gradient descent and update the weights and biases accordingly. This process represents the learning of a neural network.

<http://neuralnetworksanddeeplearning.com/chap1.html>

## 2.3 Word2Vec

### 2.3.1 Motivation

In natural language processing tasks, a good representation of words helps learning algorithms perform better. A representation is learned which maps words to vectors in a low-dimensional space compared to the vocabulary size. In this representation, we try to map context-similar words close to each other in the new vector space. We could say in an informal way: a linguistic background is made which the learning algorithm can use.

### 2.3.2 Skip-gram

One way of learning a word2vec representation of a corpus  $Text$ , is by using the skip-gram model.

Based on given words  $w$  and their contexts  $c$ , we set the parameter  $\theta$  of  $p(c|w; \theta)$  to maximize:

$$\arg \max_{\theta} \prod_{(w,c \in D)} p(c|w; \theta) \quad (2.10)$$

with  $D$  the set of all word and context pairs we extract from the corpus.

#### Finding word-context pairs

Given a sequence of words, we define their context based on n-gram. In figure 2.6, n-gram is explained on the sentence "This is a sentence".

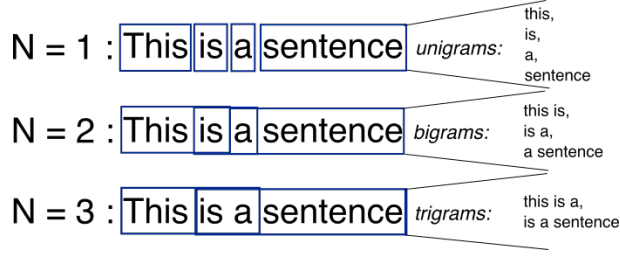


FIGURE 2.6: Explanation of n-gram.

For the Skip-gram model, we define the context of a word  $w_t$  as  $w_{t+j}$  with  $j$  between  $-c$  and  $c$ . A larger  $c$  results in more training examples and thus can lead to a higher accuracy but will have a longer training time.

#### Parameterization

We start with rewriting the conditional probability using soft-max:

$$p(c|w; \theta) = \frac{e^{v_c * v_w}}{\sum_{c' \in C} e^{v_{c'} * v_w}} \quad (2.11)$$

where  $v_c$  and  $v_w$  are vector representations for  $c$  and  $w$ , and  $C$  is the set of all available contexts. This means that the parameters  $\theta$  are  $v_{c_i}$  and  $v_{w_i}$ . Computing the optimal parameters is very expensive because you need to calculate this over all contexts  $c'$ . We also switch from product to sum by taking the logs:

$$\arg \max_{\theta} \sum_{(w,c) \in D} \log p(c|w; \theta) = \sum_{(w,c) \in D} (\log e^{v_c * v_w} - \log \sum_{c'} e^{v_{c'} * v_w}) \quad (2.12)$$



### 2.3.3 Negative Sampling

To compute the vectors using the Skip-gram model more efficiently, we introduce negative sampling.

Instead of calculating  $\sum_{c' \in C} e^{v_{c'} * v_w}$  over all contexts, we make a set  $D'$  which consists of randomly sampled word-context pairs. With this new set, we remove the costly term  $\sum_{c' \in C} e^{v_{c'} * v_w}$  and replace it with  $\sum_{(w,c) \in D'} e^{v_{c'} * v_w}$ .

In a less formal way: we are not making sure that if words appear in the same context, their vectors are more similar than all the other word vectors, but only of several vectors chosen randomly. This makes the Skip-gram model usable in terms of speed.

### 2.3.4 Neural Networks

When the word2vec algorithm is trained using the Skip-gram model, one will have a lookup table. This table contains the mapping of words to their vector representation. This lookup table can be found by training a 2-layer neural network with as goal function the function described in the previous section. The training can be done with Gradient Descent for example.

The trained 2-layer neural network can be placed in front of another neural network. It will convert the words to their vector representation and feed into the next neural network. It is empirically shown that this can improve the results of the neural network by putting the lookup table in front of it. As mentioned before, in a way, you offer background knowledge to the neural network.

<http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>

<http://arxiv.org/pdf/1402.3722v1.pdf>

<http://colah.github.io/posts/2014-07-NLP-RNNs-Representations/>

[http://homepages.inf.ed.ac.uk/ballison/pdf/lrec\\_skipgrams.pdf](http://homepages.inf.ed.ac.uk/ballison/pdf/lrec_skipgrams.pdf)

<http://recognize-speech.com/images/Antonio/Unigram.png>

<http://stackoverflow.com/questions/27860652/word2vec-negative-sampling-in-layman-term>

## 2.4 DeepWalk

DeepWalk is an approach where graph structured data is transformed into sequences of vertices. Word2vec is then applied on those sequences to learn a good vector representation for the vertices.

In figure 2.7, we see an overview of the DeepWalk algorithm. It exist of two parts.

---

**Algorithm 1** DEEPWALK( $G, w, d, \gamma, t$ )

---

**Input:** graph  $G(V, E)$   
 window size  $w$   
 embedding size  $d$   
 walks per vertex  $\gamma$   
 walk length  $t$

**Output:** matrix of vertex representations  $\Phi \in \mathbb{R}^{|V| \times d}$

- 1: Initialization: Sample  $\Phi$  from  $\mathcal{U}^{|V| \times d}$
- 2: Build a binary Tree  $T$  from  $V$
- 3: **for**  $i = 0$  to  $\gamma$  **do**
- 4:    $\mathcal{O} = \text{Shuffle}(V)$
- 5:   **for each**  $v_i \in \mathcal{O}$  **do**
- 6:      $\mathcal{W}_{v_i} = \text{RandomWalk}(G, v_i, t)$
- 7:     SkipGram( $\Phi, \mathcal{W}_{v_i}, w$ )
- 8:   **end for**
- 9: **end for**

---

FIGURE 2.7: Overview of the DeepWalk algorithm.

First a random walk generator. For each vertex  $v_i$  of the graph  $G$ , it will generate a random walk of length  $t$ . It will do this  $\gamma$  times but the order to which the vertices are traversed, is randomly ordered each pass. With those walks, a sequence of vertices is generated.

Secondly, those vertices are used for word2vec. This process is explained in section 2.3.

<http://arxiv.org/pdf/1403.6652v2.pdf>

## 2.5 Conclusion

The final section of the chapter gives an overview of the important results of this chapter. This implies that the introductory chapter and the concluding chapter don't need a conclusion.

## Chapter 3

# The Final Chapter

Morbi malesuada hendrerit dui. Nunc mauris leo, dapibus sit amet, vestibulum et, commodo id, est. Pellentesque purus. Pellentesque tristique, nunc ac pulvinar adipiscing, justo eros consequat lectus, sit amet posuere lectus neque vel augue. Cras consectetur libero ac eros. Ut eget massa. Fusce sit amet enim eleifend sem dictum auctor. In eget risus luctus wisi convallis pulvinar. Vivamus sapien risus, tempor in, viverra in, aliquet pellentesque, eros. Aliquam euismod libero a sem.

### 3.1 The First Topic of this Chapter

#### 3.1.1 Item 1

##### Sub-item 1

Nunc velit augue, scelerisque dignissim, lobortis et, aliquam in, risus. In eu eros. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Curabitur vulputate elit viverra augue. Mauris fringilla, tortor sit amet malesuada mollis, sapien mi dapibus odio, ac imperdiet ligula enim eget nisl. Quisque vitae pede a pede aliquet suscipit. Phasellus tellus pede, viverra vestibulum, gravida id, laoreet in, justo. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Integer commodo luctus lectus. Mauris justo. Duis varius eros. Sed quam. Cras lacus eros, rutrum eget, varius quis, convallis iaculis, velit. Mauris imperdiet, metus at tristique venenatis, purus neque pellentesque mauris, a ultrices elit lacus nec tortor. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent malesuada. Nam lacus lectus, auctor sit amet, malesuada vel, elementum eget, metus. Duis neque pede, facilisis eget, egestas elementum, nonummy id, neque.

##### Sub-item 2

Proin non sem. Donec nec erat. Proin libero. Aliquam viverra arcu. Donec vitae purus. Donec felis mi, semper id, scelerisque porta, sollicitudin sed, turpis. Nulla in urna. Integer varius wisi non elit. Etiam nec sem. Mauris consequat, risus nec

congue condimentum, ligula ligula suscipit urna, vitae porta odio erat quis sapien. Proin luctus leo id erat. Etiam massa metus, accumsan pellentesque, sagittis sit amet, venenatis nec, mauris. Praesent urna eros, ornare nec, vulputate eget, cursus sed, justo. Phasellus nec lorem. Nullam ligula ligula, mollis sit amet, faucibus vel, eleifend ac, dui. Aliquam erat volutpat.

#### 3.1.2 Item 2

Fusce vehicula, tortor et gravida porttitor, metus nibh congue lorem, ut tempus purus mauris a pede. Integer tincidunt orci sit amet turpis. Aenean a metus. Aliquam vestibulum lobortis felis. Donec gravida. Sed sed urna. Mauris et orci. Integer ultrices feugiat ligula. Sed dignissim nibh a massa. Donec orci dui, tempor sed, tincidunt nonummy, viverra sit amet, turpis. Quisque lobortis. Proin venenatis tortor nec wisi. Vestibulum placerat. In hac habitasse platea dictumst. Aliquam porta mi quis risus. Donec sagittis luctus diam. Nam ipsum elit, imperdiet vitae, faucibus nec, fringilla eget, leo. Etiam quis dolor in sapien porttitor imperdiet.

## 3.2 The Second Topic

Cras pretium. Nulla malesuada ipsum ut libero. Suspendisse gravida hendrerit tellus. Maecenas quis lacus. Morbi fringilla. Vestibulum odio turpis, tempor vitae, scelerisque a, dictum non, massa. Praesent erat felis, porta sit amet, condimentum sit amet, placerat et, turpis. Praesent placerat lacus a enim. Vestibulum non eros. Ut congue. Donec tristique varius tortor. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nam dictum dictum urna.

Phasellus vestibulum orci vel mauris. Fusce quam leo, adipiscing ac, pulvinar eget, molestie sit amet, erat. Sed diam. Suspendisse eros leo, tempus eget, dapibus sit amet, tempus eu, arcu. Vestibulum wisi metus, dapibus vel, luctus sit amet, condimentum quis, leo. Suspendisse molestie. Duis in ante. Ut sodales sem sit amet mauris. Suspendisse ornare pretium orci. Fusce tristique enim eget mi. Vestibulum eros elit, gravida ac, pharetra sed, lobortis in, massa. Proin at dolor. Duis accumsan accumsan pede. Nullam blandit elit in magna lacinia hendrerit. Ut nonummy luctus eros. Fusce eget tortor.

Ut sit amet magna. Cras a ligula eu urna dignissim viverra. Nullam tempor leo porta ipsum. Praesent purus. Nullam consequat. Mauris dictum sagittis dui. Vestibulum sollicitudin consectetur wisi. In sit amet diam. Nullam malesuada pharetra risus. Proin lacus arcu, eleifend sed, vehicula at, congue sit amet, sem. Sed sagittis pede a nisl. Sed tincidunt odio a pede. Sed dui. Nam eu enim. Aliquam sagittis lacus eget libero. Pellentesque diam sem, sagittis molestie, tristique et, fermentum ornare, nibh. Nulla et tellus non felis imperdiet mattis. Aliquam erat volutpat.

### 3.3 Conclusion

Vestibulum sodales ipsum id augue. Integer ipsum pede, convallis sit amet, tristique vitae, tempor ut, nunc. Nam non ligula non lorem convallis hendrerit. Maecenas hendrerit. Sed magna odio, aliquam imperdiet, porta ac, aliquet eget, mi. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Vestibulum nisl sem, dignissim vel, euismod quis, egestas ut, orci. Nunc vitae risus vel metus euismod laoreet. Cras sit amet neque a turpis lobortis auctor. Sed aliquam sem ac elit. Cras velit lectus, facilisis id, dictum sed, porta rutrum, nisl. Nam hendrerit ipsum sed augue. Nullam scelerisque hendrerit wisi. Vivamus egestas arcu sed purus. Ut ornare lectus sed eros. Suspendisse potenti. Mauris sollicitudin pede vel velit. In hac habitasse platea dictumst.

Suspendisse erat mauris, nonummy eget, pretium eget, consequat vel, justo. Pellentesque consectetur erat sed lacus. Nullam egestas nulla ac dui. Donec cursus rhoncus ipsum. Nunc et sem eu magna egestas malesuada. Vivamus dictum massa at dolor. Morbi est nulla, faucibus ac, posuere in, interdum ut, sapien. Proin consectetur pretium urna. Donec sit amet nibh nec purus dignissim mattis. Phasellus vehicula elit at lacus. Nulla facilisi. Cras ut arcu. Sed consectetur. Integer tristique elit quis felis consectetur eleifend. Cras et lectus.

Ut congue malesuada justo. Curabitur congue, felis at hendrerit faucibus, mauris lacus porttitor pede, nec aliquam turpis diam feugiat arcu. Nullam rhoncus ipsum at risus. Vestibulum a dolor sed dolor fermentum vulputate. Sed nec ipsum dapibus urna bibendum lobortis. Vestibulum elit. Nam ligula arcu, volutpat eget, lacinia eu, lobortis ac, urna. Nam mollis ultrices nulla. Cras vulputate. Suspendisse at risus at metus pulvinar malesuada. Nullam lacus. Aliquam tempus magna. Aliquam ut purus. Proin tellus.



## Chapter 4

# Conclusion

The final chapter contains the overall conclusion. It also contains suggestions for future work and industrial applications.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In

hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.



# Appendices



# Appendix A

## The First Appendix

Appendices hold useful data which is not essential to understand the work done in the master thesis. An example is a (program) source. An appendix can also have sections as well as figures and references<sup>[1]</sup>.

### A.1 More Lorem

Quisque facilisis auctor sapien. Pellentesque gravida hendrerit lectus. Mauris rutrum sodales sapien. Fusce hendrerit sem vel lorem. Integer pellentesque massa vel augue. Integer elit tortor, feugiat quis, sagittis et, ornare non, lacus. Vestibulum posuere pellentesque eros. Quisque venenatis ipsum dictum nulla. Aliquam quis quam non metus eleifend interdum. Nam eget sapien ac mauris malesuada adipiscing. Etiam eleifend neque sed quam. Nulla facilisi. Proin a ligula. Sed id dui eu nibh egestas tincidunt. Suspendisse arcu.

#### A.1.1 Lorem 15–17

Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.

Nulla mattis luctus nulla. Duis commodo velit at leo. Aliquam vulputate magna et leo. Nam vestibulum ullamcorper leo. Vestibulum condimentum rutrum mauris. Donec id mauris. Morbi molestie justo et pede. Vivamus eget turpis sed nisl cursus tempor. Curabitur mollis sapien condimentum nunc. In wisi nisl, malesuada at, dignissim sit amet, lobortis in, odio. Aenean consequat arcu a ante. Pellentesque porta elit sit amet orci. Etiam at turpis nec elit ultricies imperdiet. Nulla facilisi.

In hac habitasse platea dictumst. Suspendisse viverra aliquam risus. Nullam pede justo, molestie nonummy, scelerisque eu, facilisis vel, arcu.

Curabitur tellus magna, porttitor a, commodo a, commodo in, tortor. Donec interdum. Praesent scelerisque. Maecenas posuere sodales odio. Vivamus metus lacus, varius quis, imperdiet quis, rhoncus a, turpis. Etiam ligula arcu, elementum a, venenatis quis, sollicitudin sed, metus. Donec nunc pede, tincidunt in, venenatis vitae, faucibus vel, nibh. Pellentesque wisi. Nullam malesuada. Morbi ut tellus ut pede tincidunt porta. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam congue neque id dolor.

### A.1.2 Lorem 18–19

Donec et nisl at wisi luctus bibendum. Nam interdum tellus ac libero. Sed sem justo, laoreet vitae, fringilla at, adipiscing ut, nibh. Maecenas non sem quis tortor eleifend fermentum. Etiam id tortor ac mauris porta vulputate. Integer porta neque vitae massa. Maecenas tempus libero a libero posuere dictum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aenean quis mauris sed elit commodo placerat. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Vivamus rhoncus tincidunt libero. Etiam elementum pretium justo. Vivamus est. Morbi a tellus eget pede tristique commodo. Nulla nisl. Vestibulum sed nisl eu sapien cursus rutrum.

Nulla non mauris vitae wisi posuere convallis. Sed eu nulla nec eros scelerisque pharetra. Nullam varius. Etiam dignissim elementum metus. Vestibulum faucibus, metus sit amet mattis rhoncus, sapien dui laoreet odio, nec ultricies nibh augue a enim. Fusce in ligula. Quisque at magna et nulla commodo consequat. Proin accumsan imperdiet sem. Nunc porta. Donec feugiat mi at justo. Phasellus facilisis ipsum quis ante. In ac elit eget ipsum pharetra faucibus. Maecenas viverra nulla in massa.

## A.2 Lorem 51

Maecenas dui. Aliquam volutpat auctor lorem. Cras placerat est vitae lectus. Curabitur massa lectus, rutrum euismod, dignissim ut, dapibus a, odio. Ut eros erat, vulputate ut, interdum non, porta eu, erat. Cras fermentum, felis in porta congue, velit leo facilisis odio, vitae consectetur lorem quam vitae orci. Sed ultrices, pede eu placerat auctor, ante ligula rutrum tellus, vel posuere nibh lacus nec nibh. Maecenas laoreet dolor at enim. Donec molestie dolor nec metus. Vestibulum libero. Sed quis erat. Sed tristique. Duis pede leo, fermentum quis, consectetur eget, vulputate sit amet, erat.

## Appendix B

# The Last Appendix

Appendices are numbered with letters, but the sections and subsections use arabic numerals, as can be seen below.

### B.1 Lorem 20-24

Nulla ac nisl. Nullam urna nulla, ullamcorper in, interdum sit amet, gravida ut, risus. Aenean ac enim. In luctus. Phasellus eu quam vitae turpis viverra pellentesque. Duis feugiat felis ut enim. Phasellus pharetra, sem id porttitor sodales, magna nunc aliquet nibh, nec blandit nisl mauris at pede. Suspendisse risus risus, lobortis eget, semper at, imperdiet sit amet, quam. Quisque scelerisque dapibus nibh. Nam enim. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc ut metus. Ut metus justo, auctor at, ultrices eu, sagittis ut, purus. Aliquam aliquam.

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetur quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Etiam suscipit aliquam arcu. Aliquam sit amet est ac purus bibendum congue. Sed in eros. Morbi non orci. Pellentesque mattis lacinia elit. Fusce molestie velit in ligula. Nullam et orci vitae nibh vulputate auctor. Aliquam eget purus. Nulla auctor wisi sed ipsum. Morbi porttitor tellus ac enim. Fusce ornare. Proin ipsum enim, tincidunt in, ornare venenatis, molestie a, augue. Donec vel pede in lacus sagittis porta. Sed hendrerit ipsum quis nisl. Suspendisse quis massa ac nibh pretium cursus. Sed sodales. Nam eu neque quis pede dignissim ornare. Maecenas eu purus ac urna tincidunt congue.

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel

magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Maecenas non massa. Vestibulum pharetra nulla at lorem. Duis quis quam id lacus dapibus interdum. Nulla lorem. Donec ut ante quis dolor bibendum condimentum. Etiam egestas tortor vitae lacus. Praesent cursus. Mauris bibendum pede at elit. Morbi et felis a lectus interdum facilisis. Sed suscipit gravida turpis. Nulla at lectus. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Praesent nonummy luctus nibh. Proin turpis nunc, congue eu, egestas ut, fringilla at, tellus. In hac habitasse platea dictumst.

### B.2 Lorem 25-27

Vivamus eu tellus sed tellus consequat suscipit. Nam orci orci, malesuada id, gravida nec, ultricies vitae, erat. Donec risus turpis, luctus sit amet, interdum quis, porta sed, ipsum. Suspendisse condimentum, tortor at egestas posuere, neque metus tempor orci, et tincidunt urna nunc a purus. Sed facilisis blandit tellus. Nunc risus sem, suscipit nec, eleifend quis, cursus quis, libero. Curabitur et dolor. Sed vitae sem. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Maecenas ante. Duis ullamcorper enim. Donec tristique enim eu leo. Nullam molestie elit eu dolor. Nullam bibendum, turpis vitae tristique gravida, quam sapien tempor lectus, quis pretium tellus purus ac quam. Nulla facilisi.

Duis aliquet dui in est. Donec eget est. Nunc lectus odio, varius at, fermentum in, accumsan non, enim. Aliquam erat volutpat. Proin sit amet nulla ut eros consectetur cursus. Phasellus dapibus aliquam justo. Nunc laoreet. Donec consequat placerat magna. Duis pretium tincidunt justo. Sed sollicitudin vestibulum quam. Nam quis ligula. Vivamus at metus. Etiam imperdiet imperdiet pede. Aenean turpis. Fusce augue velit, scelerisque sollicitudin, dictum vitae, tempor et, pede. Donec wisi sapien, feugiat in, fermentum ut, sollicitudin adipiscing, metus.

Donec vel nibh ut felis consectetur laoreet. Donec pede. Sed id quam id wisi laoreet suscipit. Nulla lectus dolor, aliquam ac, fringilla eget, mollis ut, orci. In pellentesque justo in ligula. Maecenas turpis. Donec eleifend leo at felis tincidunt consequat. Aenean turpis metus, malesuada sed, condimentum sit amet, auctor a, wisi. Pellentesque sapien elit, bibendum ac, posuere et, congue eu, felis. Vestibulum mattis libero quis metus scelerisque ultrices. Sed purus.

# Bibliography

- [1] D. Adams. *The Hitchhiker's Guide to the Galaxy*. Del Rey (reprint), 1995.  
ISBN-13: 978-0345391803.

## Master thesis filing card

*Student:* Milan van der Meer

*Title:* Learning a Disease Embedding using Generalized Word2Vec Approaches.

*UDC:* 621.3

*Abstract:*

Here comes a very short abstract, containing no more than 500 words.  $\text{\LaTeX}$  commands can be used here. Blank lines (or the command  $\text{\backslash par}$ ) are not allowed!

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Thesis submitted for the degree of Master of Science in Engineering: Computer Science, specialisation Artificial Intelligence

*Thesis supervisor:* Prof. dr. Roel Wuyts

*Assessors:* Ir. Kn. Owsmuch  
K. Nowsrest

*Mentors:* Ir. An Assistent  
S. Dhooghe