

***EXPLORING TOPOLOGICAL RELATIONS IN WORD EMBEDDINGS.***

Mahesh Thapa

**EXPLORING TOPOLOGICAL RELATIONS IN WORD EMBEDDINGS**

Dissertation supervised by

Dr. Francisco Ramos

Co-supervised by

Dr. Roberto Henriques

Dr. Oscar Belmonte Fernández

Feb 2017

**Author’s Declaration**

Acknowledgments

To Professor ….

**EXPLORING TOPOLOGICAL RELATIONS IN WORD VECTORS**

# ABSTRACT

Word embedding preserve the syntactic relations between words. The synaptic relations between the words in a word embedding is quantified by transforming the high dimensional word embedding to two dimension. This research explores if the topological relations between spatial entities are also preserved in the word embedding. T-SNE is used to transform the high dimensional word embedding into two dimension. Cosine similarity between spatial words are measured which share similar topological relations to see if they can

# KEYWORDS

Word embedding

Natural Language Processing

Bag of Words

Skip Gram

Corpus

t-SNE

# ACRONYMS

**t-SNE** – t-Distributed Stochastic Neighbor Embedding

# **INDEX** OF **THE TEXT**

[ABSTRACT 5](#_Toc500353128)

[KEYWORDS 6](#_Toc500353129)

[ACRONYMS 7](#_Toc500353130)

[INDEX OF THE TEXT 8](#_Toc500353131)

[INDEX OF TABLES 9](#_Toc500353132)

[INDEX OF FIGURES 10](#_Toc500353133)

[1. INTRODUCTION 11](#_Toc500353134)

[1.1 Theoretical Framework 11](#_Toc500353135)

# INDEX OF TABLES

# INDEX OF FIGURES

## INTRODUCTION

### Introduction

### Motivation

### Aims and Objectives

## Literature Review

## sdfs

**Examples of bibliographic references**

You can use any bibliographic referencing system, provided that:

* It is relevant in the area of the dissertation;
* It is used consistently;
* all elements of the identification document (by document type) are present;
* all references removed from the Internet indicate, in addition to other factors, the URL and the date of appointment.

Here are some examples:

**1) Articles of periodicals**

Egenhoffer, M. J., e Franzosa, R. D., 1991, Point-set topological spatial relations. *International Journal of Geographical Information Systems*, **5**, 161-174.

**1.1) Articles of periodicals in press**

Egenhoffer, M. J., e Franzosa, R. D., in press, Point-set topological spatial relations. *International Journal of Geographical Information Systems*.

**1.2) Summaries of periodicals**

Medyckyj-Scott, D., Cuthbertson, M., e Newman, I., 1996, Discovering environmental data [Resumo]: metadabases, network information resource tools and the GENIE system. *International Journal of Geographical Information Systems*, **10**, 65-84.

**2) Books**

Heywood, I., Cornelius, S., e Carver, S., 2002, *An introduction to geographical information systems* (2ª ed.) (Harlow: Prentice Hall).

**2.1) Published Books**

Dykes, J.; MacEacren, A.M.; Kraak, M.-J. (Ed.s), 2005, *Exploring geovisualization* (Amsterdam: Elsevier).

**3) Articles / book chapters**

Openshaw, S., 1991, Developing appropriate spatial analysis methods for GIS. In *Geographical Information Systems: Principles and Applications*, editado por D. J. Maguire, M. F. Goodchild e D. W. Rhind (London: Longman), pp. 389-402.

**4) Articles in proceedings**

Miller, D., 1992, Analysis of vegetation succession within an expert system. In *Proceedings of the 5th International Symposium on Spatial Data Handling* (Columbia: International Geographical Union), pp. 381-400.

**4.1) Contributions not published in conferences**

Ruby, J., e Fulton, C., 1993, Beyond redlining: editing software that works. Poster presented in *5th International Symposium on Spatial Data Handling* (Columbia: International Geographical Union).

**5) Reports, PhD dissertations, Master dissertations**

Hearnshaw, H., 1991, *Mental models of spatial databases*. Research report 27, Midlands Regional Research Laboratory, Leicester.

Caldeira, P., 1999, *Influência da informação hipermédia na aprendizagem e satisfação: impacto do formato de apresentação da informação na aprendizagem e na satisfação em sujeitos com orientações contrastantes para o estudo*. PhD dissertation, ISEGI-UNL, Lisboa.

Bação, F., 1997, *Os sistemas de informação geográfica e as empresas*. Master dissertation, ISEGI-UNL, Lisboa.

**6) electronic resources**

**6.1) Websites**

Gittings, B., 1995, *Digital evaluation data catalogue* (URL: http//www.geo.ed.ac.uk/home/ded.html, consulta em 04-02-2004).

**6.1.1) Institutional websites**

ISEGI-UNL, 2003, Página do Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa (URL: www.isegi.unl.pt, consulta em 08-02-2007)

**6.2) Periodicals on Internet**

Egenhoffer, M. J., e Franzosa, R. D., 1991, Point-set topological spatial relations [Electronic version]. *International Journal of Geographical Information Systems*, **5**, 161-174.

Järvelin, K., e Wilson, T., 2003, On conceptual models for information seeking and retrieval research. *Information Research*, **9 (1)** (URL: http://informationr.net/ir/9-1/paper163.html, Retrieved 09-02-2004).

**6.3) “Books” on CD-ROM**

Clark, I., e Harper, W., 2000, *Pratical geostatistics* [CD-ROM] (Columbus, Ohio: Ecosse North America).

**6.4) Software**

Miller, M. E., 1993, The Interactive Tester (Versão 4.0) [Software] (Westminster, CA: Psytek Services).

**7) Legislation**

Reis, J. A. (Anot.), 1980, *Código do processo civil anotado* (3ª ed. reimp.) (Coimbra: Coimbra Editora).

Decreto-Lei n.º 192/89. *D.R. I Série*, **131 (08-06-1989)**, 2254-2257.

**8) Maps**

Missão Hidrográfica de Angola e São Tomé, 1958, [*Carta de São Tomé: Ribeira Afonso: levantamento aerofotogramétrico*], Scale 1:25000, 1 carta ([Lisboa]: Ministério do Ultramar).

Instituto da Água, 2001, [*Planta de condicionantes: plano de ordenamento da orla costeira: troço Alcobaça-Mafra*], Scale 1:25000, 1 carta ([Lisboa]: Instituto da Água).

**8.1) Map series**

Serviços Cartográficos do Exército, 1970, [*Carta militar de Portugal, Évora-Monte (Estremoz)*], Escala 1:25000, 1 carta ([Lisboa]: Serviços Cartográficos do Exército). (Série M 888, Folha 439).

**9) Interviews**

Painho, M., 2004, *[Título do guião da entrevista]* [Sound record, dd-mm-aaaa] (Lisboa: ISEGI - UNL).

**Note:**

The elements in square brackets, [ ], are elements “built” by the author of the bibliographic references, that is not on the information sources but are important for its proper identification.

In the case of electronic resources, information between [] seeks to explain the type of the resource.

For maps, as they often don’t have title, the title is "built" from the information available on them.



***Exploring Topological Relations in Word Embedding***

Mahesh Thapa

simbolo

Mahesh Thapa

2017

***Exploring topological relations in word embeddings***

