



TRABAJO FIN DE GRADO

GRADO EN INGENIERÍA INFORMÁTICA

Diseño e implementación de un analizador de dependencias para procesamiento de lenguaje natural en Español

Usando Máquinas de Soporte Vectoriales

Autor

Alejandro Alcalde Barros

Directores

Salvador García



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

—
6 de noviembre de 2016

Alejandro Alcalde Barros: *Diseño e implementación de un analizador de dependencias para procesamiento de lenguaje natural en Español, Usando Máquinas de Soporte Vectoriales*, Grado en Ingeniería Informática, ©
6 de noviembre de 2016

SUPERVISORS:
Salvador García
Put name here
Put name here

LOCATION:
Granada

TIME FRAME:
6 de noviembre de 2016

Ohana means family.
Family means nobody gets left behind, or forgotten.
— Lilo & Stitch

Dedicated to the loving memory of Rudolf Miede.
1939–2005

RESUMEN

En este trabajo se implementa un método para analizar dependencias palabra a palabra con una estrategia de abajo a arriba (*Bottom-Up*) usando Máquinas de Soporte Vectoriales (SVMs). En concreto, este trabajo se ha centrado en analizar las dependencias entre palabras en Castellano. Aunque la precisión de los resultados no está cerca del estado del arte, es necesario tener en cuenta que este parseador no usa información sobre la estructura de las frases.

ETIQUETAS: PNL, SVM, Parseo de dependencias.

ABSTRACT

In this project, a method for analyzing word to word dependencies is implemented using a bottom-up strategy with the help of Support Vector machines. In particular, this work has focused in analyzing dependencies between words for spanish language. Even though accuracy is far from the state of the art, it is worth noting this parser is not using information about the sentences structure.

TAGS: NLP, SVM, Dependency parsing.

ÍNDICE GENERAL

I	MOTIVACIÓN E INTRODUCCIÓN	1
II	OBJETIVOS DEL TRABAJO	3
III	RESOLUCIÓN DEL TRABAJO	5
IV	CONCLUSIONES Y VÍAS FUTURAS	7
V	APPENDIX	9
A	APPENDIX TEST	11
A.1	Appendix Section Test	11
A.2	Another Appendix Section Test	11
	BIBLIOGRAFÍA	13

ÍNDICE DE FIGURAS

ÍNDICE DE CUADROS

Cuadro 1	Autem usu id	11
----------	--------------	--------------------

LISTINGS

Listing 1	A floating example (listings manual)	11
-----------	--------------------------------------	--------------------

ACRONYMS

Parte I

MOTIVACIÓN E INTRODUCCIÓN

Entre otros elementos, y no necesariamente en este orden, la introducción debe servir para contextualizar el trabajo (explicando antecedentes importantes para el desarrollo realizado y, en su caso, realizando un estudio del estado del arte) y describir el problema entre manos, de forma que el lector tenga desde este momento una idea clara del problema a resolver o del producto a desarrollar, y una visión general de la solución alcanzada. Se concluirá esta sección con el alcance de la memoria, de tal manera que el lector pueda hacerse una idea del contenido posterior de la misma.

Las principales fuentes consultadas deben ser citadas aquí.

Parte II

OBJETIVOS DEL TRABAJO

En este apartado deberán aparecer con claridad los objetivos, inicialmente previstos en la propuesta de TFG, indicando el alcance para cada uno de ellos. Es conveniente indicar de manera precisa las interdependencias entre los distintos objetivos y también enlazarlos con los diferentes apartados de la memoria.

Los aspectos formativos previos más utilizados pueden ser destacados aquí.

Parte III

RESOLUCIÓN DEL TRABAJO

Se explicarán los métodos y procesos empleados para desarrollar el trabajo y alcanzar los objetivos. Es conveniente destacar tanto los métodos inicialmente previstos como aquellos que hayan tenido que ser agregados en el desarrollo del trabajo.

Éste es el lugar de presentar todos los datos técnicos y científicos realizados en el TFG. Debe ser detallado, claro y preciso.

En caso de ser un TFG en el que se desarrolle software, se recomienda que la sección 5 quede estructurada de la siguiente forma:

5.1 Planificación y presupuesto. Planificación temporal, con su correspondiente división en fases y tareas, y la posterior comparación con los datos reales obtenidos tras realizar el proyecto. También se incluir un presupuesto del trabajo a realizar.

5.2 Análisis y diseño. Se incluirá la especificación de requerimientos y la metodología de desarrollo por la que se ha optado, así como los “planos” del proyecto, que contendrán las historias de usuario o casos de uso, diagrama conceptual, diagramas de iteración, diagramas de diseño, esquema arquitectónico y bocetos de las interfaces de usuario. Además, se describirán las estructuras de datos fundamentales y los desarrollos algorítmicos no triviales.

5.3. Implementación y pruebas. En esta sección se incluirán todos los aspectos relacionados con la programación de la aplicación y las tecnologías seleccionadas, justificándolas e incluyendo el diseño de pruebas e informes de ejecución de las mismas.

Parte IV

CONCLUSIONES Y VÍAS FUTURAS

Las conclusiones deben incluir todas aquellas de tipo profesional y académico. Además, se debe indicar si los objetivos han sido alcanzados totalmente, parcialmente o no alcanzados.

Si hubiese claramente posibles vías de desarrollo posterior es interesante destacarlas, poniéndolas en valor en el contexto inicial del trabajo.

Parte V

APPENDIX

APPENDIX TEST

Lorem ipsum at nusquam appellantur his, ut eos erant homero concludaturque. Albucius appellantur deterruisset id eam, vivendum partiendo dissentiet ei ius. Vis melius facilisis ea, sea id convenire referrentur, takimata adolescens ex duo. Ei harum argumentum per. Eam vidit exerci appetere ad, ut vel zzril intellegam interpretaris.

More dummy text.

A.1 APPENDIX SECTION TEST

Test: [Tabla 1](#) (This reference should have a lowercase, small caps A if the option floatperchapter is activated, just as in the table itself → however, this does not work at the moment.)

LABITUR BONORUM PRI NO	QUE VISTA	HUMAN
fastidii ea ius	germano	demonstratea
suscipit instructor	titulo	personas
quaestio philosophia	facto	demonstrated

Cuadro 1: Autem usu id.

A.2 ANOTHER APPENDIX SECTION TEST

Equidem detraxit cu nam, vix eu delenit periculis. Eos ut vero constituto, no vidit propriae complectitur sea. Diceret nonummy in has, no qui eligendi recteque consetetur. Mel eu dictas suscipiantur, et sed placerat oporteat. At ipsum electram mei, ad aequae atomorum mea. There is also a useless Pascal listing below: [Listing 1](#).

Listing 1: A floating example (listings manual)

```
for i:=maxint downto 0 do
begin
{ do nothing }
end;
```


DECLARATION

Put your declaration here.

Granada, 6 de noviembre de 2016

Alejandro Alcalde Barros

COLOPHON

This document was typeset using the typographical look-and-feel classicthesis developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". classicthesis is available for both L^AT_EX and L^YX:

<https://bitbucket.org/amiede/classicthesis/>

Happy users of classicthesis usually send a real postcard to the author, a collection of postcards received so far is featured here:

<http://postcards.miede.de/>