



Tesis
Automatización de lectura de Currículum Vitae
para selección de personal

Calonge, Federico Matias
calongefederico@gmail.com

27 de julio de 2021

Resumen

En la Tesis de Ingeniería que se presenta, se diseña un *sistema de lectura automática de Curriculum Vitae*. La finalidad del mismo es ayudar al reclutador laboral a elegir a los mejores candidatos para los puestos laborales que tenga disponible mediante una medición de similitud entre textos: Curriculum Vitae de los candidatos por un lado, y keywords que referencian a los puestos laborales por el otro. El sistema estará desarrollado utilizando el lenguaje de programación Python, y permitirá verificar la teoría desarrollada.

Abstract

This Engineering Thesis introduces an *automatic Curriculum Vitae reading system*. The purpose of it is to help the job recruiter to choose the best candidates for the available job positions by means of a measurement of similarity between texts: Curriculum Vitae of the candidates on the one hand, and keywords that refer to job positions on the other. The system will be developed using Python programming language and will allow to verify the developed theory.

Índice

1. Introducción.	5
1.1. Objetivos del Proyecto.	5
1.1.1. Objetivo general.	5
1.1.2. Objetivos específicos.	5
1.2. Alcance del Proyecto.	5
1.3. Organización.	5
2. Reclutamiento laboral en IT.	5
2.1. Introducción.	5
2.2. Problemáticas.	6
2.3. Sistemas de lectura y análisis de CV.	6
3. Natural Language Processing.	6
3.1. Introducción.	6
3.2. Preprocesamiento de textos.	6
3.3. Similitud entre textos.	6
3.4. Técnicas para medir Similitud entre textos.	6
3.4.1. Cosine Similarity.	6
3.4.2. Word Mover's Distance (WMD).	7
3.5. Algoritmos de vectorización.	7
3.5.1. TF-IDF.	7
3.5.2. Word Embeddings.	7
3.5.2.1. ¿Cómo entrenar Word Embeddings?	7
4. Desarrollo.	7
4.1. Esquema del sistema.	7
4.2. Set de Datos.	7
4.2.1. Curriculum Vitae.	8
4.2.2. Descripciones Puestos Laborales.	8
4.3. Medición de similitud entre textos.	8
4.3.1. FALTA: Pasos de las pruebas que hice: "Mejorando... / Implementación de... / "Problemas al..."	8

4.4. Resultados Obtenidos.	8
4.5. Conclusiones.	8
4.6. Agregando funcionalidades.	8
4.6.1. Base de datos.	8
4.6.2. Framework Web.	9
4.6.3. Roles y Usuarios.	9
4.6.4. Manejo de los datos.	9
4.6.4.1. Modelado.	9
4.6.4.2. Filtrado.	9
4.6.4.3. Visualización.	9
4.7. Caso de Uso.	9
4.8. Limitaciones del sistema.	9
5. Próximos pasos.	10
6. Glosario.	10
7. Anexo.	10
8. Bibliografía.	10
9. Agradecimientos.	10

1 Introducción.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

1.1 Objetivos del Proyecto.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

1.1.1 Objetivo general.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

1.1.2 Objetivos específicos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

1.2 Alcance del Proyecto.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

1.3 Organización.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

2 Reclutamiento laboral en IT.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

2.1 Introducción.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

2.2 Problemáticas.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

2.3 Sistemas de lectura y análisis de CV.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3 Natural Language Processing.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.1 Introducción.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.2 Preprocesamiento de textos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.3 Similitud entre textos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.4 Técnicas para medir Similitud entre textos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.4.1 Cosine Similarity.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.4.2 Word Mover's Distance (WMD).

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.5 Algoritmos de vectorización.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.5.1 TF-IDF.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.5.2 Word Embeddings.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

3.5.2.1 ¿Cómo entrenar Word Embeddings?

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4 Desarrollo.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.1 Esquema del sistema.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.2 Set de Datos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.2.1 Curriculum Vitae.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.2.2 Descripciones Puestos Laborales.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.3 Medición de similitud entre textos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.3.1 FALTA: Pasos de las pruebas que hice: "Mejorando... / Implementación de... / "Problemas al...".

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.4 Resultados Obtenidos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.5 Conclusiones.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6 Agregando funcionalidades.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.1 Base de datos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.2 Framework Web.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.3 Roles y Usuarios.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.4 Manejo de los datos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.4.1 Modelado.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.4.2 Filtrado.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.6.4.3 Visualización.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.7 Caso de Uso.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

4.8 Limitaciones del sistema.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

5 Próximos pasos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

6 Glosario.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

7 Anexo.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

8 Bibliografía.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.

9 Agradecimientos.

Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo. Texto ejemplo.