Abstract automático mediante

técnicas de (NLP)



Manual de Usuario

iAbstract

Abstract automático mediante

técnicas de (NLP)

**M.Sc. Luis Eduardo Muñoz Guerrero**

Profesor titular

Universidad Tecnológica de Pereira

Universidad Tecnológica de Pereira

Facultad de Ingenierías

Ingeniería en Sistemas y Computación

Contenido

[TABLA DE IMÁGENES 4](#_Toc74203130)

[INTRODUCCIÓN 5](#_Toc74203131)

[OBJETIVOS 6](#_Toc74203132)

[CARACTERISTICAS DEL ENTORNO 6](#_Toc74203133)

[Requerimientos mínimos de hardware y software: 6](#_Toc74203134)

[USO DEL APLICATIVO 7](#_Toc74203135)

[VENTANA PRINCIPAL 7](#_Toc74203136)

[VENTANA ABSTRACT AUTOMATICO. 8](#_Toc74203137)

[VENTANA PROCESAMIENTO DE LENGUAJE NATURAL 9](#_Toc74203138)

[VENTANA PROCESAMIENTO DE LENGUAJE NATURAL 10](#_Toc74203139)

[VENTANA Stop Words 11](#_Toc74203140)

[VENTANA Tokenización 12](#_Toc74203141)

[VENTANA Tabla de Frecuencia 13](#_Toc74203142)

[VENTANA Tokenizar Oraciones 14](#_Toc74203143)

[VENTANA Pesos Oraciones 15](#_Toc74203144)

[VENTANA Suma Frecuencias 16](#_Toc74203145)

[VENTANA Promedio 17](#_Toc74203146)

[VENTANA Agrupación 18](#_Toc74203147)

[VENTANA Sorteo de Oraciones 19](#_Toc74203148)

[VENTANA Abstract 20](#_Toc74203149)

# TABLA DE IMÁGENES

[Ilustración 1 Ejecucion del programa 8](#_Toc74203215)

[Ilustración 2 Menu Abstract 9](#_Toc74203216)

[Ilustración 3 Documentación Algoritmo 10](#_Toc74203217)

[Ilustración 4 Botones Pasos 11](#_Toc74203218)

[Ilustración 5 Stop Words 12](#_Toc74203219)

[Ilustración 6 Tokenización 13](#_Toc74203220)

[Ilustración 7 Tabla de Frecuencia 14](#_Toc74203221)

[Ilustración 8 Tokenizar Oraciones 15](#_Toc74203222)

[Ilustración 9 Pesos Oraciones 16](#_Toc74203223)

[Ilustración 10 Suma Frecuencias 17](#_Toc74203224)

[Ilustración 11 Promedio 18](#_Toc74203225)

[Ilustración 12 Agrupación 19](#_Toc74203226)

[Ilustración 13 Sorteo de Oraciones 20](#_Toc74203227)

[Ilustración 14 Abstract 21](#_Toc74203228)

# INTRODUCCIÓN

Se plantea el desarrollo de un software capaz de realizar el abstract(resumen) de un artículo de investigación a fin de ayudar al autor mediante técnicas de inteligencia artificial, con el propósito de realizar un resumen detallado y especifico para que el lector pueda comprender fácilmente el tema del artículo.

El software cuenta con un algoritmo de (PNL) Programación Neurolingüística o también conocido como Procesamiento de Lenguaje Natural (PLN) a tal fin de que la maquina pueda interactuar con el texto, creando diversas aplicaciones en el ámbito de procesamiento de textos.

Los algoritmos de (PNL) se usan esencialmente para la resolución de anagramas, ChatBots, conversión de texto a voz, entre otras aplicaciones, pero en este caso y la mas importante para el software desarrollado es la posibilidad de crear resúmenes a partir de texto u audio, en este caso texto.

El desarrollo de los algoritmos son bastante simples y complejos a su vez, debido a que si se toman cantidades inmensas de texto por eficiente que sea el algoritmo todo depende de la maquina para el procesamiento del resumen. El algoritmo consiste en tokenizar el texto, lo cual quiere decir separar todo el texto primero en palabras y después en oraciones, no obstante a diferencia de la tokenizacion con las palabras, con las oraciones es un poco mas complejo y diferente, a pesar de que a ambos se les entrega un peso y una ponderación, en el caso de las oraciones es mas especial debido a que se deben aplicar algoritmos de detección de sentimientos a fin de crear oraciones coherentes al momento de unirlas en un párrafo.

# OBJETIVOS

* Aplicar métodos de inteligencia artificial para el desarrollo de artículos científicos.
* Construir abstracts a partir de algoritmos inteligentes.
* Construcción de resumen detallado del artículo, a fin de una mejor comprensión del tema a tratar

# CARACTERISTICAS DEL ENTORNO

## Requerimientos mínimos de hardware y software:

* Monitor con resolución de 1024 x 768 o superior.
* Sistema operativo Windows o versiones superiores.
* Procesador de 1.6GHz o superior.
* Memoria RAM de 1gGb o superior
* 500Mb disponibles en el disco duro.
* Solo archivos en formato PDF.
* Instalación previa o tener instalado Python en su última versión

<https://www.python.org/downloads/> (Python Download Page)

# USO DEL APLICATIVO

El programa posee múltiples vistas las cuales permiten diferentes funcionalidades. Se explicarán su uso y función en los siguientes apartados:

## VENTANA PRINCIPAL

* Al iniciar el programa se presentará la pantalla principal titulada “iAbstract”. Esta ventana cuenta con dos botones “Abstract” encargado de redirigir a la ventana encargada de realizar el resumen y el botón “Algoritmo” el cual tiene como propósito redirigir a la ventana encargada de explicar el algoritmo usado con los cuales el usuario podrá interactuar.

Se cuenta con una corta introducción a lo que es el programa.



Ilustración Ejecucion del programa

## VENTANA ABSTRACT AUTOMATICO.

* + Al ingresar a la ventana de abstracto automático se contará con un botón en la barra superior para regresar al inicio tal cual como dice su nombre; botón “inicio”.

La ventana cuenta con tres opciones:

* + - La primera opción es el botón “Generar Abstract” se encarga de cargar el artículo en formato pdf y crear el resumen de este.
    - La segunda opción es el botón “Exportar Español” la cual se encarga de exportar el resumen en formato de texto plano, es decir archivos con extensión txt, legible por cualquier dispositivo.
    - La tercera opción es el botón “Exportar Ingles” la cual se encarga de exportar el resumen en idioma ingles en formato de texto plano, es decir archivos con extensión txt, legible por cualquier dispositivo.

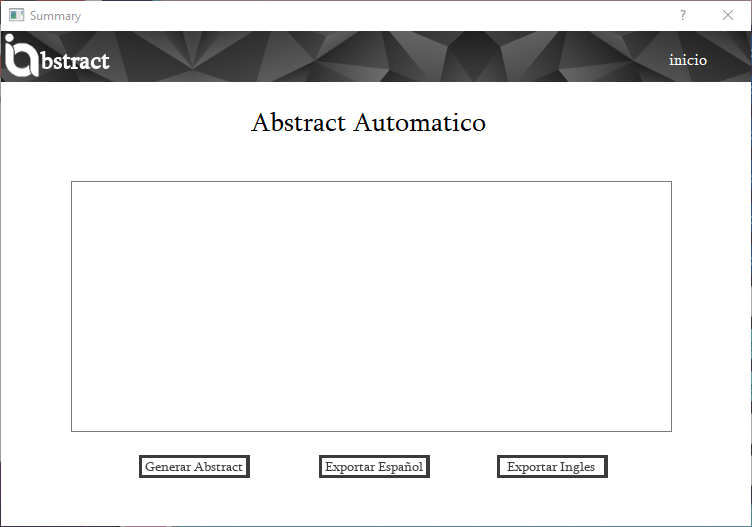


Ilustración Menu Abstract

## VENTANA PROCESAMIENTO DE LENGUAJE NATURAL

* + En la sección de “Procesamiento de lenguaje natural“, se podrá encontrar de forma teórica la explicación del algoritmo usado para realizar el resumen del artículo.
  + Al inicio de la ventana se encontrará una breve descripción del algoritmo usado
  + En la barra superior se encontrará el botón “inicio” encargado de navegar por las opciones del software.
  + La ventana cuenta con una sección de pasos donde se explica paso a paso el algoritmo

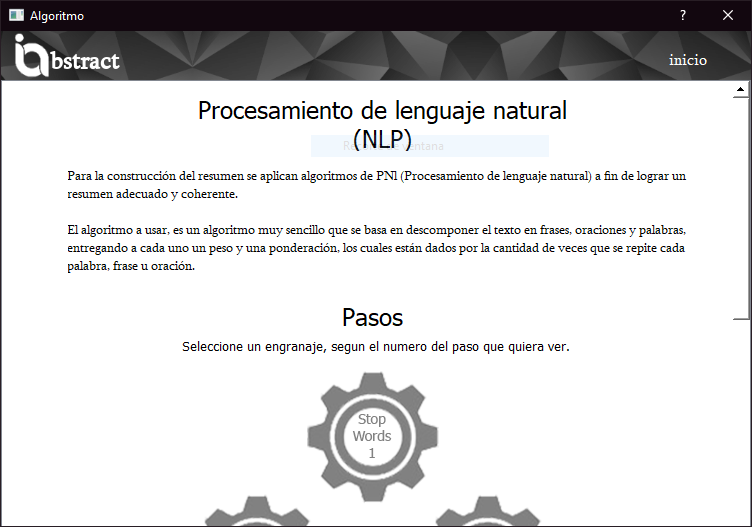


Ilustración Documentación Algoritmo

## VENTANA PROCESAMIENTO DE LENGUAJE NATURAL

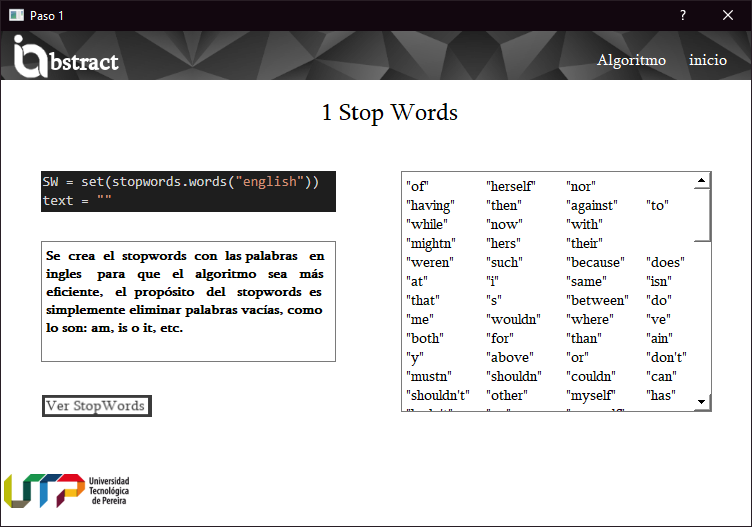
* En el apartado y sección de pasos se encontrarán imágenes con forma de engranajes, al presionar sobre un engranaje la vista cambiara por completo según el engranaje presionado.
* En total se encontrá 10 botones con los nombres y enumeración de cada paso



Ilustración Botones Pasos

## VENTANA Stop Words

* La ventana Stop Words cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

Ilustración 5 Stop Words

## VENTANA Tokenización

* La ventana Tokenización cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.



Ilustración Tokenización

## VENTANA Tabla de Frecuencia

* La ventana Tabla de Frecuencia cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

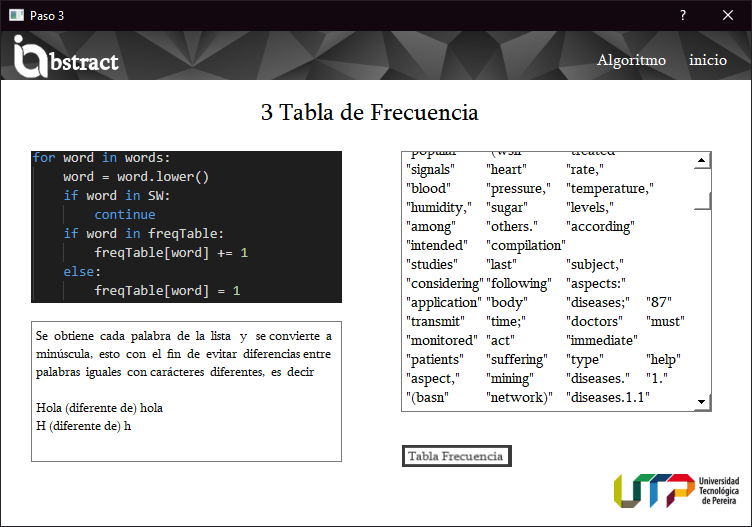


Ilustración Tabla de Frecuencia

## VENTANA Tokenizar Oraciones

* La ventana Tokenizar Oraciones cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

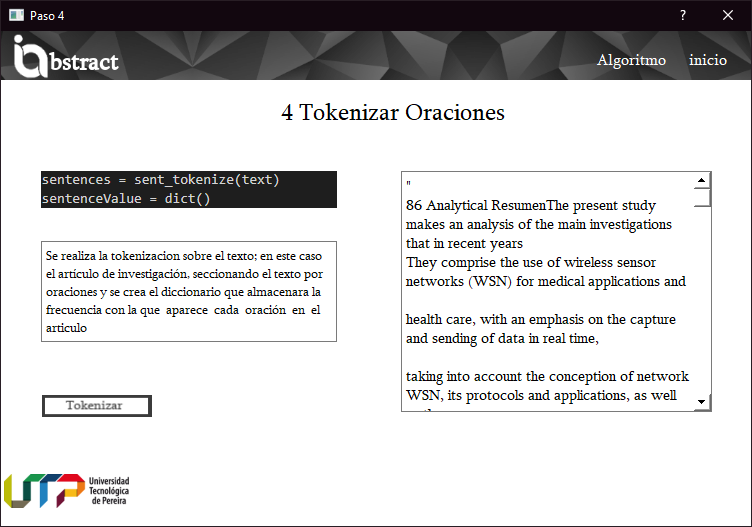


Ilustración Tokenizar Oraciones

## VENTANA Pesos Oraciones

* La ventana Pesos Oraciones cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

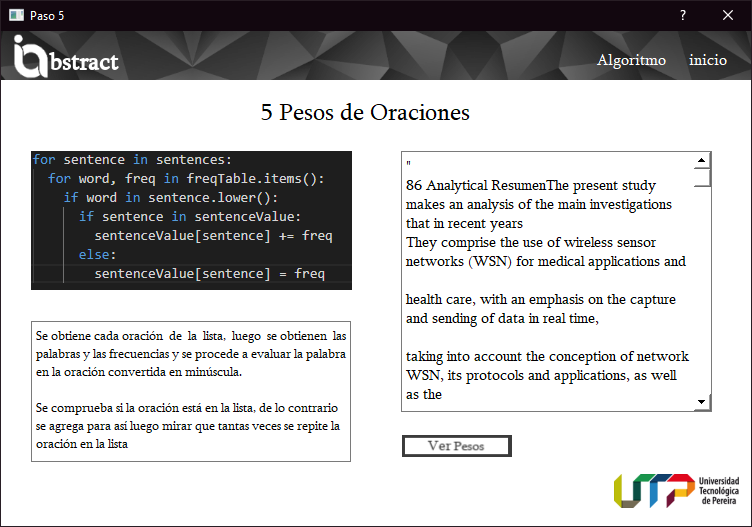


Ilustración Pesos Oraciones

## VENTANA Suma Frecuencias

* La ventana Suma Frecuencias cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

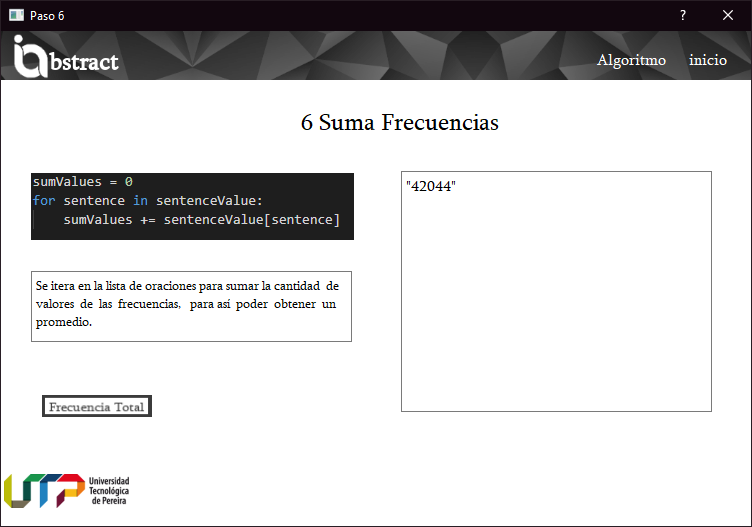


Ilustración Suma Frecuencias

## VENTANA Promedio

* La ventana Promedio cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.



Ilustración Promedio

## VENTANA Agrupación

* La ventana Agrupación cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

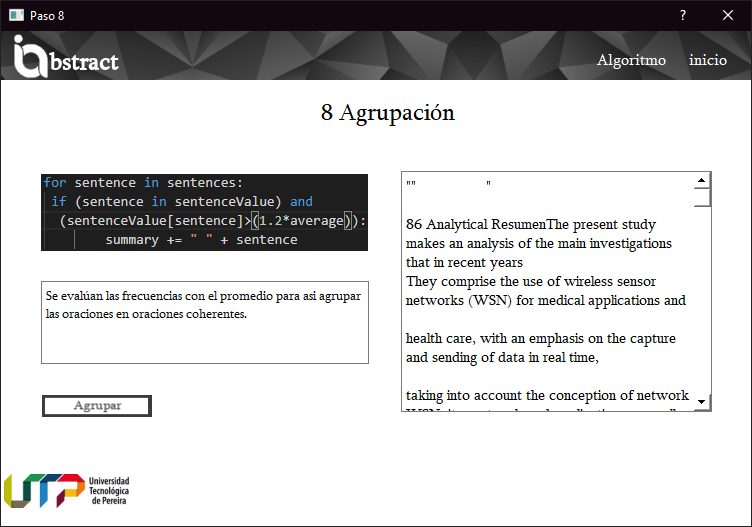


Ilustración Agrupación

## VENTANA Sorteo de Oraciones

* La ventana Sorteo de Oraciones cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

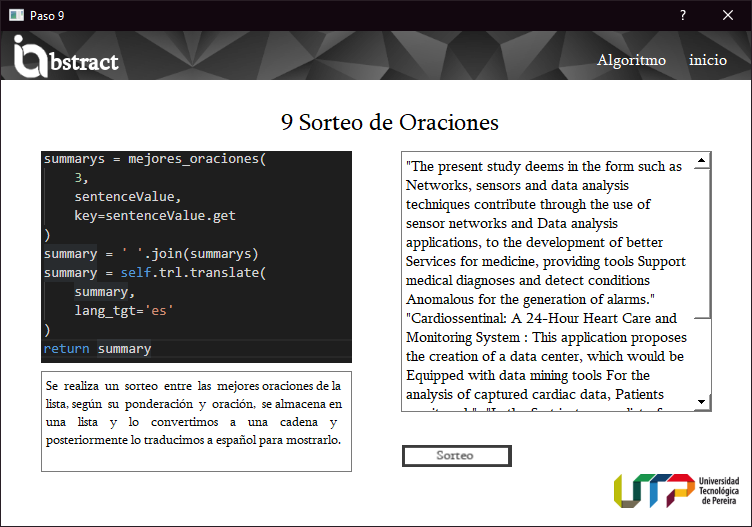


Ilustración Sorteo de Oraciones

## VENTANA Abstract

* La ventana Abstract cuenta con tres secciones, la primera es donde se muestra el paso del código, la segunda se muestra la explicación en palabras del paso y la tercera es donde se muestra el resultado del paso.
* La ventana cuenta con un botón el cual sirve para ejecutar la línea de código correspondiente al paso.

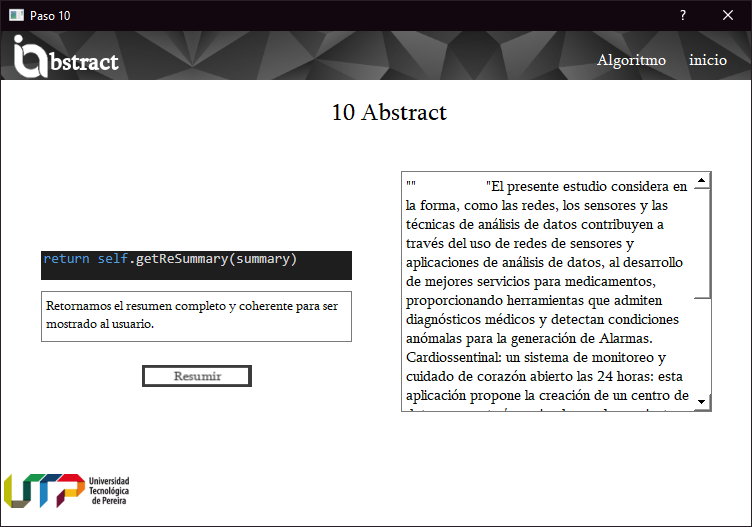


Ilustración Abstract