



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación



El futuro
es de todos

DNP
Departamento
Nacional de Planeación




Estudio de comparación del CONPES de reactivación económica con los Planes de Desarrollo Territorial (PDT) 2020-2023

Unidad de Científicos de Datos
Dirección de Desarrollo Digital

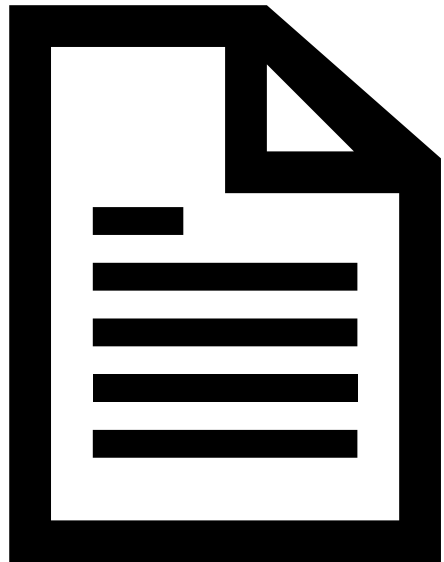
2020



Identificación de la inclusión del CONPES de reactivación económica en los PDT

<h2>¿Qué?</h2>	<p>Comparar de forma automatizada el texto del CONPES de reactivación económica con los Planes de Desarrollo Territoriales 2020-2023</p>	
<h2>¿Para qué?</h2>	<p>Para realizar un análisis de la vocación de inversión de las Entidades Territoriales,</p>	
<h2>Avances</h2>	<ol style="list-style-type: none"> 1. Primer flujo de extracción de textos, preprocesamiento, vectorización y comparación 2. Se encontraron términos clave exclusivos y relevantes dentro del CONPES de reactivación 	

Descripción del proyecto

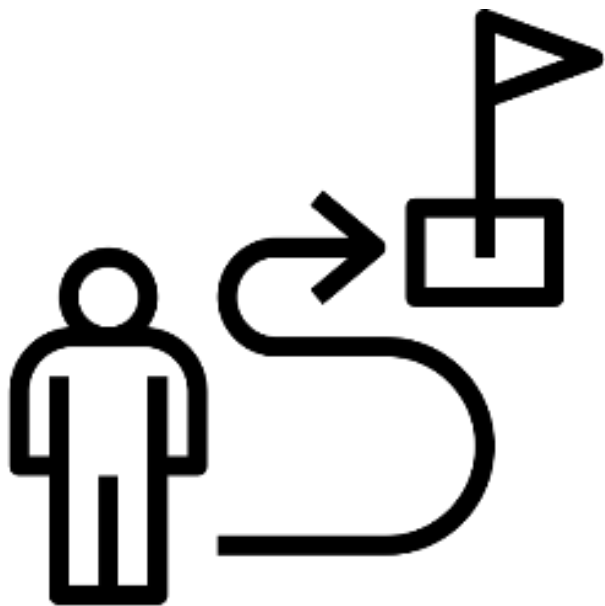


CONPES de reactivación **VS** Cientos de PDT

Automatización con Python



Metodología



- 1 Extracción de texto de PDF
- 2 Limpieza y preparación del texto
- 3 Vectorización del texto
- 4 Comparación de texto de CONPES de reactivación con PDT

Términos clave

Exclusivos al CONPES de reactivación y relevantes para ser buscados en los PDT

Exclusivos

- 1 COVID
- 2 ECNT
- 3 Bioeconomía
- 4 Govtech

Relevantes

- 1 Reactivación
- 2 Reactivación económica
- 3 Pobreza
- 4 Empresas

Vectorización y comparación

1

Representación
numérica de textos

a Bag of words

b TF-IDF

c word2vec

2

Comparación de
textos

a Similitud coseno

b Distancia euclideana

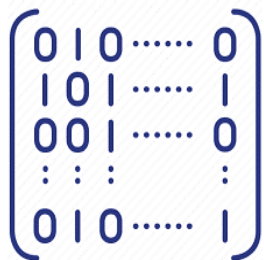
c Distancia Manhattan

d Distancia Hamming



Vectorización y comparación de textos

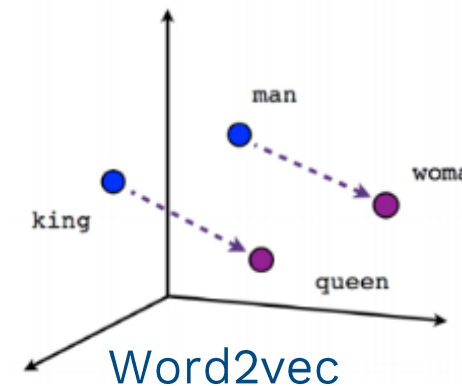
Vectorizar



TF-IDF



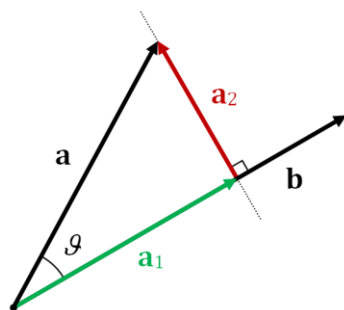
Bag of Words



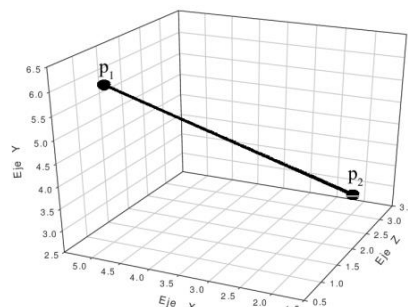
Word2vec



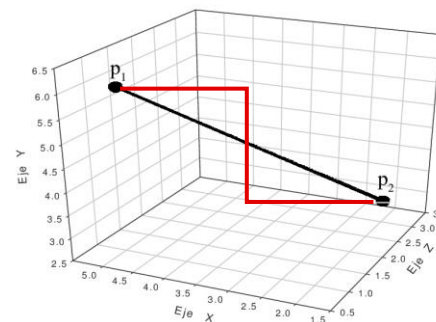
Comparar



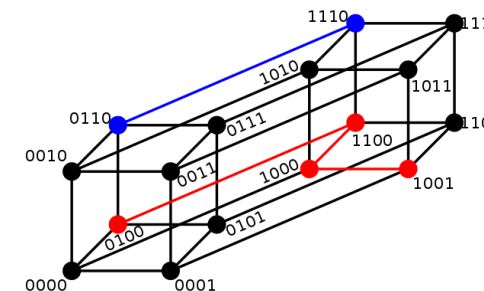
Similitud coseno



Distancia Euclidiana



Distancia Manhattan



Distancia Hamming

Siguientes pasos

1

Entregar lista de términos clave a DDDR (mañana)

2

Comparar el CONPES de reactivación con los PDT

3

Consolidar resultados en un archivo de Excel





**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación