



# NLP > APLICADO A PROYECTOS DE LEY

PROCESAMIENTO DE LENGUAJE NATURAL

Profesor: Jorge Castillo

Estudiante: Gustavo Lobos

## Introducción:

Según las estimaciones de la industria, sólo el 21% de los datos disponibles están presentes en forma estructurada. Los datos se están generando mientras hablamos, mientras twiteamos, mientras enviamos mensajes en Whatsapp y en varias otras actividades. La mayoría de estos datos existen en forma textual, que es de naturaleza altamente desestructurada. Entonces, viendo la relevancia que está ganando el Procesamiento del Lenguaje Natural, es que se ha realizado el presente trabajo de investigación sobre una de las tantas aplicaciones que tiene NLP. En este caso nos referimos a los algoritmos que existen detrás de las técnicas de "Text Summarization" o lo que son los Resúmenes de Texto. En esta oportunidad utilizaremos una biblioteca llamada Sumy.

Sumy es una biblioteca de Python para extraer resúmenes de páginas HTML o textos planos. Fue desarrollada por Miso-Belica, un estudiante Ruso/Eslovaco que desarrolló Sumy como proyecto de [Tesis de grado](#). Esta [biblioteca](#) es útil para ver como funcionan los diferentes métodos de "summarization" que se pueden listar de la siguiente manera:

## Listado de métodos de reducción/resumen de textos

1. Luhn
2. Edmundson
3. **Latent Semantic Analysis | LSA**
4. **LexRank**
5. TextRank
6. SumBasic

En esta oportunidad vamos a comparar el funcionamiento de 2 de estos métodos de resumen, que son LSA y LexRank y los vamos a aplicar a Proyectos de Ley.

## Métodologías:

A la hora de resumir texto, existen dos enfoques que deben ser considerados:

1. Técnicas Extractivas
2. Técnicas Abstractivas

### **Enfoque por Extracción:**

Es el más simple de los enfoques, que trata las fuentes como un conjunto de frases. Para llevar a cabo el proceso de extracción, el primer paso es identificar las frases más relevantes de las fuentes. Para proceder en esta fase se pueden utilizar criterios de selección basados en métodos estadísticos y otros tipos de heurísticos durante el tratamiento de la fuente, también puede resultar útil tener en cuenta las colisiones entre diferentes fuentes como indicativo de la bondad de la frase. En segundo lugar y a partir de los fragmentos extraídos previamente, hay que generar un resumen procurando perder la menor cantidad de información posible y evitando la redundancia. Por último, pueden aplicarse tratamientos posteriores al resumen para conseguir un texto coherente y bien cohesionado, uniendo los diferentes fragmentos o completando frases que puedan haber quedado incompletas.

### **Enfoque por Abstracción:**

El método abstractivo produce un resumen con palabras, frases y oraciones nuevas e innovadoras. Los resúmenes por abstracción suponen un reto mucho mayor que el de los resúmenes por extracción, ya que las técnicas de tratamiento del lenguaje natural empleadas en estos son mucho más sofisticadas. Esto es debido a que los resúmenes por abstracción no se limitan a extraer fragmentos relevantes de un texto, sino que analizan el texto con mayor profundidad para poder comprenderlo y poder generar un nuevo resumen a partir de la información analizada de la fuente.

Si la idea es comprender el texto de forma más profunda, esto ya hace que el problema sea más difícil, a esto se le añade la dificultad de generar texto desde cero a partir de una base de conocimiento. El resumen por abstracción es por lo tanto similar al método que emplearía un ser humano para hacer un resumen propio de un texto, y la dificultad de reproducir ese método por un ordenador es lo que hace que los métodos de extracción sean los que han conseguido avances mas significativos.

### **Enfoque seleccionado:**

En este notebook sólo se abordan algoritmos basados en enfoque de extracción, dado que para usar métodos abstractivos se tendrían que utilizar soluciones basadas en Machine Learning, lo que se escapa del target de esta investigación.

## **Algoritmos**

En esta oportunidad, sólo nos concentraremos en la técnica de extracción utilizando las dos variables antes mencionadas. La razón es que para el ejemplo que utilizaremos son las que mejores resultados arrojaron al momento de ejecutar el código.

1. **Análisis de la Semántica Latente | LSA:** Es una técnica que utiliza procesos estadísticos para extraer y representar los significados de las palabras. Los significados son representados en términos de su semejanza con otras palabras en

un extenso texto llamado "corpus de documentos". La ventaja práctica del LSA sobre otros tipos de mediciones es que va más allá de las similitudes léxicas tales como silla/sillas o correr/corrió, y en el mismo sentido es capaz de evaluar la similitud semántica entre términos tales como silla/mesa, mesa/madera y madera/bosque. De este modo, el Análisis Semántico Latente no solo indica que dos términos están relacionados, sino que también señala cuán similares son. El LSA comienza procesando un texto de grandes dimensiones. Este texto contiene miles de párrafos y frases y es conocido como corpus lingüístico. El corpus se representa en una matriz cuyas filas contiene todos los términos distintos del corpus (palabras). De este modo, la matriz contiene sencillamente el número de veces que cada término aparece en un documento. Sobre esta matriz de frecuencias se efectúa una ponderación con el objeto de restar importancia a las palabras excesivamente frecuentes y aumentarla a las palabras moderadamente infrecuentes. La razón de esta ponderación es sencilla: las palabras demasiado frecuentes no sirven para discriminar bien la información importante del párrafo y las moderadamente infrecuentes sí. El siguiente paso es someter a esta matriz ponderada a un algoritmo llamado Descomposición en Valores Singulares. SVD es una técnica de reducción de dimensiones.

2. **LexRank:** Es un sistema de resúmenes automático no supervisado basado en grafos y el cual está inspirado en técnicas de PageRank, o sea: Ranking Léxico. Se basa en la idea de que la importancia de una frase depende de la similitud que esta tenga con el resto de oraciones del texto. Se basa en los cálculos del coseno para calcular la similitud entre dos frases. Dos frases están conectadas si el coseno de la similitud entre ambas es superior a cierto umbral. Para la selección de las frases que forman parte del resumen, busca cuáles son los nodos —oraciones— más centrales del texto; es por ello que existen diversas medidas de centralidad las cuales pasaremos a comentar: Centralidad basada en el grado: Donde la importancia de una frase depende del número de conexiones que esta tenga con el resto de nodos. Centralidad basada en vectores propios: Pondera cada una de las conexiones con la importancia del nodo con el que se conecta; por ello un nodo será más importante cuanto mayor sea el valor de los nodos a los que se conecta. Entonces: este método tiende a seleccionar oraciones que tengan más términos centrales en común con los documentos y todos los rasgos se consideran de igual manera.

## Código

Descomentando el las líneas del siguiente código se podrán instalar las bibliotecas. Si ya están instaladas, favor de comentar nuevamente y recargar el Kernel.

```
In [ ]: #pip install sumy
        #pip install newspaper3k
```

```
In [2]: import sumy
        from sumy.parsers.plaintext import PlaintextParser
        from sumy.nlp.tokenizers import Tokenizer
        from sumy.summarizers.lex_rank import LexRankSummarizer
```

```
from sumy.summarizers.lsa import LsaSummarizer
from newspaper import Article
```

## Contexto

Para las siguientes partes de código, vamos a utilizar un poco de Web Scraping para extraer el contenido de la ley publicada en el diario oficial del Congreso Colombiano.

Para este ejemplo se utiliza el Congreso de este país dado que su publicación es a través de una página scrapeable ya que es html puro, no así el caso de Chile en donde la Biblioteca Nacional y el Diario Oficial enmascaran con Javascript.

```
In [3]: #Caso de Estudio 1
        ley1 = "https://www.suin-juriscol.gov.co/viewDocument.asp?id=30039047"
```

```
In [4]: # Descargar el articulo y convertirlo a texto plano en un gran string
        contenido1 = Article(ley1)
        contenido1.download()
        contenido1.parse()

        # articulo
        print(contenido1.text)
```

Subtipo: DECRETO LEGISLATIVO

En ejercicio de las atribuciones que le confieren el artículo 215 de la Constitución Política, en concordancia con la Ley 137 de 1994, y en desarrollo de lo previsto desarrollo de lo previsto en el Decreto 417 del 17 de marzo de 2020, «Por el cual se declara un Estado de Emergencia Económica, Social y Ecológica en todo el territorio nacional», y

CONSIDERANDO

Que, en los términos del artículo 215 de la Constitución Política, el presidente de la República, con la firma de todos los ministros, en caso de que sobrevengan hechos distintos de los previstos en los artículos 212 y 213 de la Constitución Política, que perturben o amenacen perturbar en forma grave e inminente el orden económico, social y ecológico del país, o que constituyan grave calamidad pública, podrá declarar el Estado de Emergencia Económica, Social y Ecológica.

Que, según la misma norma constitucional, una vez declarado el estado de Emergencia Económica, Social y Ecológica el Presidente, con la firma de todos los ministros, podrá dictar decretos con fuerza de ley destinados exclusivamente a conjurar la crisis y a impedir la extensión de sus efectos.

Que estos decretos deberán referirse a materias que tengan relación directa y específica con el estado de Emergencia Económica, Social y Ecológica, y podrán, en forma transitoria establecer nuevos tributos o modificar los existentes.

Que mediante el Decreto 417 del 17 de marzo de 2020 se declaró el estado de emergencia económica, social y ecológica en todo el territorio nacional por el término de treinta (30) días, con el fin de conjurar la grave calamidad pública que afecta al país por causa del nuevo Coronavirus COVID-19.

Que dentro de las razones generales tenidas en cuenta para la adopción de dicha medida se incluyeron las siguientes:

Que el 7 de enero de 2020 la Organización Mundial de la Salud identificó el nuevo Coronavirus - COVID-19 y declaró este brote como emergencia de salud pública de importancia internacional.

Que el 6 de marzo de 2020 el Ministerio de Salud y de la Protección Social dio a conocer el primer caso de brote de enfermedad por Coronavirus - COVID-19 en el territorio nacional.

Que el 11 de marzo de 2020 la Organización Mundial de la Salud - OMS declaró el actual brote de enfermedad por Coronavirus - COVID-19 como una pandemia, esencialmente por la velocidad de su propagación y la escala de transmisión, toda vez que al 11 de marzo de 2020 a la OMS se habían notificado cerca de 125.000 casos de contagio en 118 países y que a lo largo de esas últimas dos semanas el número de casos notificados fuera de la República Popular China se había multiplicado en 13 veces, mientras que el número de países afectados se había triplicado, por lo que instó a los países a tomar acciones urgentes.

Que según la OMS la pandemia del nuevo Coronavirus COVID-19 es una emergencia sanitaria y social mundial, que requiere una acción efectiva e inmediata de los gobiernos, las personas y las empresas.

Que mediante la Resolución 380 del 10 de marzo de 2020, el Ministerio de Salud y Protección Social adoptó, entre otras, medidas preventivas sanitarias de aislamiento y cuarentena de las personas que, a partir de la entrada en vigencia de la precitada resolución, arribaran a Colombia desde la República Popular China, Francia, Italia y España.

Que mediante Resolución 385 del 12 de marzo de 2020, el ministro de Salud y Protección Social, de acuerdo con lo establecido en el artículo 69 de la Ley 1753 de 2015, declaró el estado de emergencia sanitaria por causa del nuevo Coronavirus COVID-19 en todo el territorio nacional hasta el 30 de mayo de 2020 y, en virtud de la misma, adoptó una serie de medidas con el objeto de prevenir y controlar la propagación del COVID-19 y mitigar sus efectos.

Que el vertiginoso escalamiento del brote de nuevo Coronavirus COVID-19, hasta configurar una pandemia, representa actualmente una amenaza global a la salud pública, con afectaciones al sistema económico, de magnitudes impredecibles e incalculables, de la cual Colombia no podrá estar exenta.

Que mediante el Decreto 417 del 17 de marzo de 2020 se declaró el estado de emergencia económica, social y ecológica en todo el territorio nacional por el término de treinta (30) días, con el fin de conjurar la grave calamidad pública que afecta al país por causa del nuevo Coronavirus COVID-19. Dentro de las medidas a adoptarse se incluyeron las siguientes:

"Que con el propósito de limitar las posibilidades de propagación del nuevo virus Covid 19 y de proteger la salud del público en general y de los servidores públicos que los atienden, se hace necesario expedir normas de orden legal que flexibilicen la obligación de atención personalizada al usuario y se permita incluso la suspensión de términos legales en las actuaciones administrativas y jurisdiccionales.

Que con igual propósito de limitar las posibilidades de propagación del nuevo virus Covid 19 y de proteger la salud del público en general y de los servidores públicos que los atienden, se hace necesario expedir normas que habiliten actuaciones judiciales y administrativas mediante la utilización de medios tecnológicos, y adoptar las medidas pertinentes con el objeto de garantizar la prestación del servicios público de justicia, de notariado y registro, de defensa jurídica del Estado y la atención en salud en el sistema penitenciario y carcelario.

Que con el fin de evitar la propagación de la pandemia del coronavirus y contener la misma, el gobierno nacional podrá expedir normas para simplificar el proceso administrativo sancionatorio contenido en la Ley 9 de 1979 y en la Ley 1437 de 2011 garantizando el debido proceso y el derecho a la defensa.

Que ante el surgimiento de la mencionada pandemia se debe garantizar la prestación continua y efectiva de los servicios públicos, razón por la cual se deberá analizar medidas necesarias para cumplir con los mandatos que le ha entregado el ordenamiento jurídico colombiano. Lo anterior supone la posibilidad flexibilizar los criterios de calidad, continuidad y eficiencia de los servicios, establecer el orden de atención prioritaria en el abastecimiento de los mismos, flexibilizar el régimen laboral en cuanto a los requisitos de los trabajadores a contratar, implementar medidas de importación y comercialización de combustibles con el fin de no afectar el abastecimiento."

Que el artículo 3 del Decreto 417 del 17 de marzo de 2020 resolvió adoptar "mediante decretos legislativos, además de las medidas anunciadas en la parte considerativa de este decreto, todas aquellas medidas adicionales necesarias para conjurar la crisis e impedir la extensión de sus efectos, así mismo dispondrá las operaciones presupuestales necesarias para llevarlas a cabo".

Que el artículo 44 de la Constitución establece que la educación es un derecho fundamental de los niños, niñas y adolescentes, y que corresponde al Estado garantizar el adecuado cubrimiento del servicio y asegurar a los menores las condiciones necesarias para su acceso y permanencia en el sistema educativo.

Que de conformidad con la Ley 137 de 1994 "Por la cual se reglamentan los Estados de Excepción en Colombia", el Gobierno nacional debe propender por la adopción de medidas destinadas a conjurar la crisis e impedir la extensión de sus efectos, y en ningún momento, se podrá suspender el derecho a la educación, por lo que se requiere emprender acciones que permitan la continuidad de la prestación del servicio, así como del complemento alimentario que facilite el desarrollo del proceso pedagógico y de aprendizaje desde los hogares.

Que mediante Circular 020 del 16 de marzo de 2020, expedida por la Ministra de Educación Nacional, dirigida a gobernadores, alcaldes y secretarios de educación de Entidades Territoriales Certificadas en Educación, en aplicación de lo dispuesto en los numerales 5.1 y 5.2 del artículo 148 de la Ley 115 de 1994, el artículo 5 de la Ley 715 de 2001, y los artículos 2.4.3.4.1. y 2.4.3.4.2 del Decreto 1075 de 2015, Único Reglamentario del Sector Administrativo de Educación Nacional, ordenó a las secretarías de educación en todo el territorio nacional ajustar el calendario académico de Educación Preescolar, Básica y Media, para retomar el trabajo académico a partir del 20 de abril de 2020.

Que con ocasión de la declaratoria de Emergencia Sanitaria por causa del Coronavirus COVID-19 y la declaratoria de Emergencia Económica, Social y Ecológica contenida en el Decreto 417 de 2020, se hace necesario contar con herramientas que permitan la protección del derecho fundamental de educación de niños, niñas y adolescentes, en su componente de alimentación.

Que de conformidad con la Sentencia T-457 del 27 de noviembre de 2018 de la Sala Quinta de Revisión de la honorable Corte Constitucional: "la alimentación escolar es una garantía de acceso y permanencia de los niños, niñas y adolescentes en el sistema educativo, reconocida en el marco jurídico colombiano; (ii) uno de sus principales objetivos consiste en garantizar la asistencia a las aulas en condiciones dignas, sin que los estudiantes se vean expuestos al hambre y la desnutrición y, por ende, se evite la deserción escolar; igualmente, contribuye al crecimiento y desarrollo físico y psicológico adecuado; propende por el nivel de salud más alto posible; potencia la atención de los menores de edad para el aprendizaje y aumenta la matrícula escolar

Que el artículo 19 de la Ley 1176 de 2007, "Por la cual se desarrollan los artículos 356 y 357 de la Constitución Política y se dictan otras disposiciones", limita la localización y cobertura del Programa de Alimentación Escolar a los establecimientos educativos, lo cual impide que los niños, niñas, adolescentes y jóvenes del país, que con ocasión de la pandemia

del Coronavirus CQVID-19 se encuentran en sus casas, puedan consumir la alimentación escolar en sus hogares durante el receso estudiantil originado por la pandemia del Coronavirus COVID-19.

Que, por lo anterior, se requiere modificar el marco legal del Programa de Alimentación Escolar, con el fin de garantizar el derecho fundamental a la educación de los niños, niñas, adolescentes y jóvenes del país, en su componente de alimentación.

Que el sistema educativo oficial brinda atención en educación preescolar, básica y media a 6.928.742 niños, niñas y adolescentes, quienes con ocasión de las medidas adoptadas para prevenir la propagación del Coronavirus CQVID-19, deberán ser atendidos a través de estrategias pedagógicas flexibles coordinadas con las 96 Secretarías de Educación certificadas.

Que el numeral 16.3. del artículo 16 de la Ley 715 de 2001 "por la cual se dictan normas orgánicas en materia de recursos y competencias de conformidad con los artículos 151, 288, 356 y 357 (Acto Legislativo 01 de 2001) de la Constitución Política y se dictan otras disposiciones para organizar la prestación de los servicios de educación y salud, entre otros" dispone que el criterio de Equidad del Sistema General de Participaciones para educación, solo se podrá distribuir a cada distrito o municipio, una suma residual de acuerdo con el indicador de pobreza certificado por el DANE.

Que el inciso 4 del artículo 17 de la Ley 715 de 2001 señala que "Los recursos de calidad serán girados directamente a los municipios y no podrán ser utilizados para gastos de personal de cualquier naturaleza", norma que sólo permite el giro de estos recursos a los municipios.

Que los recursos asignados en virtud de los numerales 16.3 del artículo 16 y el inciso 4 del artículo 17 de la Ley 715 de 2001, son destinados, entre otros usos permitidos por la Ley, para el apalancamiento del Programa de Alimentación Escolar.

Que los departamentos, como entidades territoriales certificadas en educación, son responsables de la contratación y operación del Programa de Alimentación Escolar con cobertura para los municipios no certificados de su jurisdicción, y por lo tanto, se hace necesario distribuir y girar también a los departamentos, recursos del criterio de Equidad y Calidad que corresponden a los municipios no certificados de su jurisdicción, durante la Emergencia Económica, Social y Ecológica, con el fin de garantizar el derecho fundamental a la educación de los niños, niñas, adolescentes y jóvenes del país, en su componente de alimentación.

Que la población a la que van dirigidos estos programas y estrategias corresponde a cerca de 6,9 millones de niños, niñas y adolescentes matriculados en el sector oficial, de los cuales cerca de 3,7 millones de niños se encuentran en entidades territoriales no certificadas en educación, por lo cual recursos por el orden de \$180 mil millones del Sistema General de Participaciones se hacen necesarios para la atención de la población ubicada en las entidades territoriales no certificadas en los programas ya mencionados, los cuales deben ser distribuidos y ejecutados por los departamentos.

Que, en consecuencia, resulta necesaria la adopción de una medida de orden legislativo, tendiente a permitir que: (i) el Programa de Alimentación Escolar se brinde a los niños, niñas y adolescentes matriculados en el sector oficial para su consumo en casa, durante la vigencia del estado de Emergencia Económica, Social y Ecológica; y, (ii) además de municipios y distritos, habilitar a los departamentos recursos de los criterios de Equidad y Calidad del Sistema General de Participaciones, para garantizar la continuidad del derecho a la educación de niños, niñas y adolescentes, durante el receso estudiantil a causa de la pandemia Coronavirus COVID-19.

Que con base en lo anterior, se expidió el Decreto 470 del 24 de marzo de 2020 "Por el cual se dictan medidas que brindan herramientas a las entidades territoriales para garantizar la ejecución del Programa de Alimentación Escolar y la prestación del servicio público de educación preescolar,

básica y media, dentro del Estado de Emergencia Económica, Social y Ecológica”.

Que las medidas adoptadas en el precitado Decreto 470 del 24 de marzo 2020, se sujetaron a la vigencia del Estado de Emergencia Económica Social y Ecológica, declarada mediante Decreto 417 del 17 de marzo de 2020, cuyo artículo 1 establece el Estado de Emergencia por el término de 30 días calendario, emergencia que va hasta el 15 de abril de 2020.

Que el Ministerio de Salud y Protección Social reportó el 9 de marzo de 2020, 0 muertes y 3 casos confirmados en Colombia.

Que, pese a las medidas adoptadas, el Ministerio de Salud y Protección Social, a las 21:00 horas del 23 de marzo de 2020 reportó tres (3) muertes y 306 casos confirmados en Colombia, distribuidos así: Bogotá D.C. (114), Cundinamarca (11), Antioquia (40), Valle del Cauca (31), Bolívar (16), Atlántico (7), Magdalena (2), Cesar (1), Norte de Santander (8), Santander (3), Cauca (2), Caldas (6), Risaralda (13), Quindío (6), Huila (10), Tolima (4), Meta (1), Casanare (1), San Andrés y Providencia (1).

Que pese a las medidas adoptadas, el Ministerio de Salud y Protección Social reportó el 7 de abril de 2020 50 muertes y 1.780 casos confirmados en Colombia, distribuidos así: Bogotá D.C. (861), Cundinamarca (60), Antioquia (209), Valle del Cauca (250), Bolívar (63), Atlántico (63), Magdalena (14), Cesar (17), Norte de Santander (26), Santander (14), Cauca (14), Caldas (16), Risaralda (44), Quindío (34), Huila (34), Tolima (15), Meta (14), Casanare (2), San Andrés y Providencia (2), Nariño (7), Boyacá (13), Córdoba (7), Sucre (1) y La Guajira (1).

Que según la Organización Mundial de Salud - OMS, en reporte de fecha 23 de marzo de 2020 a las 15:51 GMT-5, se encuentran confirmados 334,981 casos, 14,652 fallecidos y 190 países, áreas o territorios con casos del nuevo coronavirus COVID- 19.

Que según la Organización Mundial de la Salud - OMS, en reporte de fecha 7 de abril de 2020 a las 19:00 GMT-5, - Hora del Meridiano de Greenwich-, se encuentran confirmados 1,282,931 casos, 72,774 fallecidos y 211 países, áreas o territorios con casos del nuevo coronavirus COVID-19.

Que debido al aumento de casos de contagio y muertes del Coronavirus COVID-19 en el país, reportado por el Ministerio de Salud y Protección Social, la señora ministra de Educación Nacional, mediante Directiva 09 del 7 de abril de 2020, extendió hasta el 31 de mayo de 2020, el periodo de aislamiento preventivo en los establecimientos educativos de preescolar, básica y media.

Que el señor presidente de la República, mediante Decreto 531 del 8 de abril de 2020 “Por el cual se imparten instrucciones en virtud de la emergencia sanitaria generada por la pandemia del Coronavirus COVID-19, y el mantenimiento del orden público”, ordenó el aislamiento preventivo obligatorio de todas las personas habitantes de la República de Colombia, a partir de las cero horas (00:00 a.m.) del día 13 de abril de 2020, hasta las cero horas (00:00 a.m.) del día 27 de abril de 2020.

Que con el objetivo de seguir contribuyendo a las medidas de aislamiento preventivo obligatorio, en pro de la salud pública y el bienestar general de las niñas, niños, jóvenes y la población más vulnerable, resulta necesario adoptar medidas que garanticen el derecho a la educación en su componente de alimentación, mientras estos dan continuidad de las jornadas de trabajo académico en casa, a partir del 20 de abril y hasta cuando se determine, durante la emergencia sanitaria, según Directiva 09 del 7 de abril de 2020 del Ministerio de Educación Nacional.

En mérito de lo expuesto,

DECRETA



```
#Vamos a tokenizar nuestro texto utilizando nuestro lenguaje natural
parser = PlaintextParser.from_string(contenido1.text,Tokenizer("spanish"))
parser
```

Out[5]: <sumy.parsers.plaintext.PlaintextParser at 0x285bebfed60>

```
In [6]: # Usando LexRank
lexrank = LexRankSummarizer()
# Usando LSA
lsa = LsaSummarizer()
#Resumir el proyecto de ley en 3 ideas
resumen_lexrank = lexrank(parser.document, 3)
resumen_lsa = lsa(parser.document, 3)
```

```
In [7]: start = "\033[1m"
end = "\033[0;0m"
```

```
In [8]: print(start + 'Resumen utilizando algoritmo LexRank:' + end)
print()
for ideas in resumen_lexrank:
    print(ideas, '\n')
```

#### **Resumen utilizando algoritmo LexRank:**

Que mediante Resolución 385 del 12 de marzo de 2020, el ministro de Salud y Protección Social, de acuerdo con lo establecido en el artículo 69 de la Ley 1753 de 2015, declaró el estado de emergencia sanitaria por causa del nuevo Coronavirus COVID-19 en todo el territorio nacional hasta el 30 de mayo de 2020 y, en virtud de la misma, adoptó una serie de medidas con el objeto de prevenir y controlar la propagación del COVID-19 y mitigar sus efectos.

Que, en consecuencia, resulta necesaria la adopción de una medida de orden legislativo, tendiente a permitir que: (i) el Programa de Alimentación Escolar se brinde a los niños, niñas y adolescentes matriculados en el sector oficial para su consumo en casa, durante la vigencia del estado de Emergencia Económica, Social y Ecológica; y, (ii) además de municipios y distritos, habilitar a los departamentos recursos de los criterios de Equidad y Calidad del Sistema General de Participaciones, para garantizar la continuidad del derecho a la educación de niños, niñas y adolescentes, durante el receso estudiantil a causa de la pandemia Coronavirus COVID-19.

Que con base en lo anterior, se expidió el Decreto 470 del 24 de marzo de 2020 "Por el cual se dictan medidas que brindan herramientas a las entidades territoriales para garantizar la ejecución del Programa de Alimentación Escolar y la prestación del servicio público de educación preescolar, básica y media, dentro del Estado de Emergencia Económica, Social y Ecológica".

```
In [9]: print(start + 'Resumen utilizando algoritmo LSA:' + end)
print()
for ideas in resumen_lsa:
    print(ideas, '\n')
```

#### **Resumen utilizando algoritmo LSA:**

Que según la OMS la pandemia del nuevo Coronavirus COVID-19 es una emergencia sanitaria y social mundial, que requiere una acción efectiva e inmediata de los gobiernos, las personas y las empresas.

Que el sistema educativo oficial brinda atención en educación preescolar, básica y media a 6.928.742 niños, niñas y adolescentes, quienes con ocasión de las medidas adoptadas para prevenir la propagación del Coronavirus CQVID-19, deberán ser atendidos a través de estrategias pedagógicas flexibles coordinadas con las 96 Secretarías de Educación certificadas.

(861), Cundinamarca (60), Antioquia (209), Valle del Cauca (250), Bolívar (63), Atlántico (63), Magdalena (14), Cesar (17), Norte de Santander (26), Santander (14), Cauca (14), Caldas (16), Risaralda (44), Quindío (34), Huila (34), Tolima (15), Meta (14), Casanare (2), San Andrés y Providencia (2), Nariño (7), Boyacá (13), Córdoba (7), Sucre (1) y La Guajira (1).

## Conclusiones

Como vimos anteriormente en como funciona en grandes términos cada algoritmo, vemos que a pesar de que LSA pareciera ser la solución predilecta por estos sistemas de resúmenes, cuando están aplicados a textos de una complejidad superior como es el caso de las leyes, estos tienen a perder un poco la objetividad del contenido, aún en caso de LSA que es producto del análisis semántico de las palabras. A esto se le debe sumar que en este caso no hicimos reducción por stopwords dado que es en estos enfoques en donde ganan relevancia. Resumir textos sin stopwords no tendría sentido dado que perdería la cohesión y coherencia.

Como vimos, el Natural Language Processing es una especialidad muy compleja: en ella, los ordenadores tienen que hacer frente a casos concretos y, en el caso de las palabras ambiguas, existe la probabilidad de que el ordenador las interprete erróneamente. Y el resultado también puede verse afectado por el método que utilicemos para una tarea determinada.

- Botana, G. de J. and Cascón, J. A. L. (no date) 'Una aproximación distribuida al análisis semántico.', p. 447.
- Revista signos: estudios de lingüística, ISSN 0035-0451, ISSN-e 0718-0934, N°. 59, 2005, págs. 303-323
- Hernández Osuna, S. and Ferreira Cabrera, A. (2016) 'Procesamiento semántico automático, enfocado en la coherencia textual, para apoyar la producción escrita de noticias', Estudios filológicos, (58), pp. 97-122. doi: 10.4067/S0071-17132016000200005.

## BIBLIOGRAFÍA

Gracias!

In [ ]: