

Métodos estadísticos de machine learning aplicados en el estudio de la accesibilidad web: una revisión de la literatura

Zambrano Félix

<https://orcid.org/0000-0003-1383-658X>

felix.zambrano@utm.edu.ec

Maestrante del Instituto Posgrado de la
Universidad Técnica de Manabí
Portoviejo-Ecuador

Muñoz Emanuel

<https://orcid.org/0000-0002-0997-0578>

emanuel.munoz@utm.edu.ec

Universidad Técnica de Manabí
Portoviejo-Ecuador

Recibido (10/09/2022), Aceptado (06/04/2023)

Resumen: Una web accesible se refiere a que cualquier persona sobre todo las que tengan alguna discapacidad física puedan acceder al contenido del website sin problemas. El objetivo de esta investigación es el análisis mediante revisiones bibliográficas de los métodos de machine learning aplicados al estudio de la accesibilidad en los portales web de los Gobiernos Autónomos Descentralizados. Además, se utilizó una metodología de revisión sistemática de literatura a más de veinte artículos científicos relacionados con palabras clave como: accesibilidad web, estadística, machine learning, entre otros. En los resultados obtenidos se destacan varias técnicas, sobre todo, las del aprendizaje no supervisado ya que se observó su utilidad en varias investigaciones, mejorando el análisis y comprensión de los datos. Esta investigación ha demostrado que se pueden hacer trabajos interesantes de accesibilidad web en las instituciones, teniendo en cuenta que estos estudios serían un aporte significativo para mejorar el acceso a los contenidos.

Palabras clave: Accesibilidad web, machine learning, aprendizaje no supervisado, estadística.

Statistical machine learning methods applied in the study of web accessibility: a literature review

Abstract. - Technological development in solid-state chemistry, nanotechnology, and new materials is advancing at an accelerated pace; studies of methods to generate thin films of conductive and semiconductor materials are of great interest; however, current methods tend to be very expensive and inaccessible to developing countries. This work seeks to present viable and economical alternatives for teaching laboratories to investigate chemical deposition processes in aqueous solutions and produce thin layers of materials of interest in solid-state chemistry and new materials. The metals studied were copper, cobalt and, nickel in different salts and reducing agents, hydrazine hydrochloride, phenylhydrazine and, sodium borohydride. The main results show that it is possible to use cheaper chemicals to study depositions in an aqueous solution, a viable alternative for laboratories.

Keywords: Web accessibility, machine learning, unsupervised learning, statistics.



I. INTRODUCCIÓN

Vivimos en una época en donde la tecnología tiene una importancia vital y el concepto de accesibilidad debe de ir ligado a todo lo que se desarrolla en base a las tecnologías de la información y comunicación (TIC). Con la creación del internet, también se desarrollaron los portales web con el objetivo de que las personas naveguen y se comuniquen. Esta tecnología ha evolucionado al punto de que hoy en día tener un sitio web es la presentación al mundo de una empresa o institución. Sin embargo, es importante destacar que no basta con tener una web atractiva, hay que diseñarla de forma que sea accesible al mayor número de personas posible, independientemente de su condición física, más aún si se trata de un portal en línea que pretende simplificar los trámites municipales que todo ciudadano debe de realizar, y estos sitios web en el caso de Ecuador, los tienen los Gobiernos Autónomos Descentralizados (GAD).

Según la Organización Mundial de la Salud (OMS), más de mil millones de personas, es decir, casi el 15% de la población mundial, sufre algún tipo de discapacidad. El número de personas con discapacidad va en aumento, lo que se explica por la creciente prevalencia de dolencias crónicas y el envejecimiento de la población. En este sentido se han hecho muchas investigaciones acerca de la accesibilidad y en cada una de ellas se han implementado diversas técnicas estadísticas para recolectar, ordenar, clasificar y analizar los datos. Estos estudios se han realizado a diferentes tipos de portales web como por ejemplo a los del área de salud, en el área del sector turístico, en la educación y también se han hecho estudios a los websites de instituciones gubernamentales. Además, esta clase de investigación que se han realizado en el Ecuador en su gran mayoría por lo general están dirigidos a las instituciones de educación superior y hay muy poca información acerca de estudios realizados a instituciones públicas gubernamentales. Es por eso que el objetivo de esta revisión literaria es descubrir y entender como se ha llevado acabo otras investigaciones aplicadas a la accesibilidad y que métodos estadísticos relacionados a machine learning son las más adecuados para el tratamiento de los datos.

En este trabajo, la recolección de la información se realiza mediante los motores de búsqueda en bases de datos especializados en artículos científicos, se seleccionaron veinte documentos que aportan significativamente a este trabajo, estos pasaron por un proceso de revisión y análisis con el fin de que el contenido sea de calidad y tenga un gran aporte para el desarrollo de esta investigación.

II. DESARROLLO

En esta sección abarcaremos los aspectos teóricos y conceptuales relacionados a la accesibilidad web y a los métodos de machine learning que se aplican a este tipo de estudio.

A. Accesibilidad en sitios web

El significado de accesibilidad está relacionado con el concepto de diseño universal, ya que nos dice que el diseño accesible debe ser utilizado por la mayor cantidad de personas sin restricciones, independientemente de sus habilidades físicas o intelectuales, con el objetivo de facilitar a cualquiera el acceso al contenido de un documento electrónico, incluso al usuario que tenga una discapacidad física [1].

La accesibilidad tiene como finalidad permitir que muchos individuos usen los sitios web, independientemente de los conocimientos, las habilidades personales y las características técnicas del dispositivo, eliminando las barreras que no permite que los adultos mayores y las personas con discapacidad utilicen esta tecnología. De esta manera se puede afirmar que, utilizando correctamente los criterios de accesibilidad, un portal web se convierte en una herramienta muy útil porque facilita el acceso a la información a un mayor número de personas, especialmente aquellas que tienen alguna limitación física [2].

B. Estándares y normas de accesibilidad

Las normas más utilizadas por los investigadores cuando realizan este tipo de estudio son las sugeridas por la WAI (Web Accessibility Initiative), que es una rama de la W3C (World Wide Consortium) que se dedica a mejorar la accesibilidad de la web [3]. La WAI tiene una pauta de accesibilidad llamada WCAG 2.1 (Web Content Accessibility Guidelines, Pautas de Accesibilidad del Contenido Web), que nos da especificaciones de cómo debería ser el contenido que se publica en las páginas web como videos, textos, imágenes, enlaces, entre otros [4]. De tal manera, dentro de la WCAG 2.1 encontramos cuatro principios que se describen como: [5].

- **Perceptible:** El contenido como la información y componentes que conforman la interfaz de usuario se deben presentar a los usuarios de forma clara de modo que puedan percibirlos.
- **Operable:** Los componentes de la interfaz de usuario deben de ser operable.
- **Comprensible:** La información que contiene la página web y el manejo de la interfaz de usuario debe ser comprensible.
- **Robusto:** El contenido de la página web debe ser lo suficientemente estable y robusto para que se adapte y sea interpretado de forma confiable por diferentes agentes de usuarios, incluyendo las ayudas técnicas.

Cada principio está formado por pautas que son los que se deben de tomar en cuenta para que una página web sea accesible. Las pautas deben de estar dentro de los criterios de conformidad para comprobar si realmente se cumplen. En este sentido, los criterios de conformidad tienen tres niveles el A, AA, AAA, en donde el AA es el que más se debe de tener en cuenta ya que en este nivel se podrá alcanzar una accesibilidad importante dentro de los sitios web [6].

C. Métodos y técnicas estadísticas aplicados al estudio de la accesibilidad web

Los métodos y técnicas estadísticas que se aplican para el tratamiento y análisis de los datos cuando se trata de temas relacionados a la web son varios. Por ejemplo, en relación a la accesibilidad las técnicas utilizadas por los investigadores son el análisis factorial de componentes principales [7], las técnicas para medir la simetría de los datos, aplican la inferencia estadística, realizan pruebas de normalidad mostrándolos en un gráficos histogramas y diagramas de cajas [2].

Otros autores [8], utilizan una técnica de clasificación para determinar cuál es el nivel de accesibilidad. Sin embargo, otras investigaciones [9], aplican técnicas estadísticas univariadas, bivariadas y multivariadas como el análisis de componentes principales, de conglomerados y factoriales de correspondencias simples. También se han aplicado técnicas de machine learning para el análisis de la accesibilidad en las presentaciones [10], donde detallan un nuevo enfoque para valorar los documentos accesibles aplicando el aprendizaje automático con un modelo construido a partir de las características de la apariencia de una presentación.

D. Machine Learning (Aprendizaje automático)

El aprendizaje automático o machine learning nació en el campo de la inteligencia artificial, incluye un conjunto de métodos matemáticos y estadísticos, cuyas tareas están relacionadas con el reconocimiento, diagnóstico, predicción, entre otros. En este sentido, existen categorías de aprendizaje, de los cuales, los más utilizados para el estudio de la web son el supervisado y no supervisado. Algunas investigaciones [11], proponen un método de aprendizaje semisupervisado para evaluar la accesibilidad de todas las páginas de un sitio web. Otros métodos [11] son conocidos como predicción activa. Este procedimiento convierte la evaluación de la accesibilidad web en un modelo de aprendizaje activo para luego hacer predicciones, logrando alcanzar una alta precisión en las evaluaciones.

El aprendizaje supervisado permite deducir patrones e hipótesis a través de los datos que se suministran al modelo, lo cual lleva a predecir una instancia correspondiente a los datos de entrada. Además, dentro del aprendizaje supervisado se pueden aplicar varias técnicas como la de árbol de decisión [12], red bayesiana y máquina de vectores de apoyo (SVM) para poder clasificar cada uno de los conjuntos de datos de los sitios web etiquetados como accesibles y no accesibles. Sin embargo, los resultados con estas técnicas pueden ser prometedores, por lo que algunos autores sugieren una clasificación automatizada de las páginas web con respecto a la accesibilidad.

Así como hay estudios aplicando el aprendizaje supervisado también lo hay utilizando el aprendizaje no supervisado [13], donde se aplica la técnica de agrupamiento clúster a 21 portales web, con lo cual plantean una técnica estadística para extraer características de las URL de forma no supervisada. La finalidad es agrupar esas particularidades para construir patrones que representen los diferentes tipos de direcciones web de un sitio.

El aprendizaje no supervisado puede liberarnos de la necesidad de etiquetar los datos y también de aplicar características de ingeniería manualmente, gracias a los métodos flexibles, generales y automatizados del aprendizaje automático [14]. También podemos encontrar un diagrama de las técnicas utilizados en los aprendizajes no supervisados, en donde los investigadores dividen las técnicas en seis categorías principales: aprendizaje jerárquico, agrupación de datos, modelos de variables latentes, reducción de dimensionalidad y detección de valores atípicos.

A continuación, se analizan algunas técnicas del aprendizaje no supervisado que han sido aplicadas a estudios relacionados a portales web, cabe recalcar que solo se revisaron técnicas que aportaron un valor significativo para la accesibilidad web.

Agrupación de datos (data clustering)

La agrupación en clústeres es un método de aprendizaje no supervisado que tiene como objetivo encontrar patrones ocultos en datos de entrada no etiquetados en forma de grupos, es decir, abarca la disposición de los datos en agrupaciones naturales significativas sobre la base de la similitud [14]. Las técnicas de agrupamiento y de clasificación son muy útiles en estos tipos de estudios porque se puede partir de unos pocos ejemplos de entrenamiento etiquetados automáticamente permitiendo captar características valiosas de los objetos de estudios [15].

Algunas investigaciones muestran que las metodologías generales permiten el agrupamiento basado en modelos, que además proporcionan un enfoque estadístico apoyados en los principios de la accesibilidad web, por lo tanto, este tipo de modelos puede ser útil para otros problemas en el análisis multivariante, en el análisis discriminante y la estimación de densidad multivariante [16]. También existen estudios enfocados en la agrupación basados en la estructura y el estilo del sitio web para el proceso de categorización, limpieza, detección de esquemas y extracciones automáticas de los datos [17]. En este sentido las técnicas de clustering aplican un enfoque de aprendizaje automático para agrupar según las barreras detectadas y seleccionando páginas representativas, se pueden obtener buenos resultados y así confirmar la validez de los niveles de accesibilidad de los portales estudiados [18].

Reducción de dimensiones

La reducción de dimensionalidad de los datos es una tarea no supervisada, donde en lugar de elegir un subconjunto de características, crea nuevas características (dimensiones) como una función de todas las características. Es útil para modelado, compresión y visualización de datos [14].

Una de las técnicas utilizadas para la reducción de dimensiones es la del análisis factorial, aunque también forma parte de la categoría de modelos de variables latente. Es diferente de otros modelos de variables latentes en términos de la variación tolerada para diferentes dimensiones. En el modelo de análisis factorial, las variables latentes tienen covarianza diagonal en lugar de covarianza isotrópica. Además del análisis factorial también se puede aplicar la técnica del análisis de componentes principales (PCA). El PCA es una técnica estadística que utiliza la transformación ortogonal en los datos para convertir n número de variables posiblemente correlacionadas en menor número k de variables no correlacionadas denominadas componentes principales. Usando esta técnica, podemos reconocer las características más fuertes que tiene el conjunto de datos, lo que hace que los datos sean más fáciles para explorar y visualizar.

En esta revisión de la literatura se encontraron algunas investigaciones en donde se aplicaron las técnicas del análisis factorial y el análisis de componentes principales. Los investigadores al aplicar estos métodos pueden dividir las características de los sitios web en componentes considerándolos como subindicadores y con la rotación varimax se puede facilitar la explicación de los vínculos entre variables y componentes [7]. Incluso combinando las técnicas de clustering y el PCA se obtiene un modelo con valores relevantes para alcanzar los objetivos afines a la accesibilidad del contenido digital [19].

E. Leyes de la Accesibilidad web en el Ecuador

En el Ecuador existe el Instituto Ecuatoriano de Normalización, INEN, que fue creado para establecer normas técnicas ecuatorianas para satisfacer los sectores productivos y servicios. El INEN en su Reglamento Técnico Ecuatoriano RTE-INEN: RTE-288 en la resolución N° 16 008 por el Ministerio de Industria y Productividad "Accesibilidad para el contenido web", establece las normas técnicas para la web, que toda institución pública y privada que brinde algún servicio a través de la web debería cumplir [20]. Este Reglamento Técnico Ecuatoriano aplica para los contenidos publicados en los portales web tanto del sector público y privado que brinden servicios por este medio, garantizando el acceso a la información y comunicación de todas las personas con y sin discapacidad. En este sentido, se ha tomado como referencia la NTE INEN-ISO/IEC 40500, Tecnología de la información-Directrices de accesibilidad para el contenido web del W3C (WCAG 2.0). Los requisitos primordiales para que un sitio web sea accesible según el reglamento del INEN son los siguientes:

El contenido publicado y todo el sitio web, tiene que cumplir con las pautas y criterios de conformidad establecidos en la Norma NTE INEN-ISO/IEC 40500, vigente. Se pide que los sitios web cumplan por completo del nivel de conformidad AA, establecido en la Norma NTE INEN-ISO/IEC 40500, vigente.

III. METODOLOGÍA

La búsqueda bibliográfica de los artículos citados en este documento se la realiza mediante los sitios web Google Scholar, Scielo, Research Gate, Science Direct, entre otros. Las palabras clave utilizadas para la búsqueda fueron "accesibilidad web", "clúster", "estadística", "aprendizaje automático", "machine learning", "aprendizaje no supervisado" entre otras.

Se encontraron más de cincuenta artículos que contenían una o más de las palabras clave, luego se escogieron los que aplicaban los métodos y técnicas del aprendizaje automático o machine learning para el tratamiento de los datos, de los cuales los seleccionados para este trabajo fueron revisados y analizados cuidadosamente, quedando finalmente veinte, desechando los otros ya que no se ajustaban a esta investigación. Además, también se considera la relevancia en los resultados obtenidos, las conclusiones y trabajos futuros.

IV. RESULTADOS

La accesibilidad de los sitios web es cada vez más necesaria ya que aporta con herramientas de utilidad a una gran diversidad de personas, que deben incorporarse a las actividades online de diferente índole, de ahí que resulte de gran interés la gestión apropiada de sitios web accesibles.

Los modelos de aprendizajes automáticos basados en las técnicas de machine learning permiten enriquecer en gran medida el estudio de la accesibilidad de los contenidos en los portales web, sin embargo, hay que saber escoger cuál técnica o método aplicar y va a depender del objetivo que se quiera alcanzar y del tipo de datos que se recolecte de los objetos de análisis. Además, se debe tener en consideración las características que están normalizadas para que una web cumpla con los requisitos mínimos y sea de fácil acceso.

La diferencia más importante entre las técnicas del aprendizaje supervisado y el aprendizaje no supervisado es que el primero, necesita que los datos estén etiquetados, es decir que la información sea identificada o procesada previamente, pero en el caso del aprendizaje no supervisado no hace falta que los datos sean etiquetados, lo que significa que no es necesario procesarlos previos a la aplicación de la técnica.

La técnica de agrupamiento o clustering es una de las más utilizadas en las investigaciones de la accesibilidad web, porque facilita la clasificación de acuerdo a características relevantes que tienen los individuos de estudio. Además, si va acompañado con la aplicación de otra técnica como la del PCA que permite comprender cuáles son las variables más relevantes que influyen en el proceso de agrupación, se logra mejorar el desarrollo y la obtención de los resultados.

Otra de las técnicas que destaca en estos tipos de investigaciones es la de reducción de dimensiones que ayudan a entender cuál es la correlación de las variables de estudio y aplicando modelos basados en agentes inteligentes se logra comprender los niveles de accesibilidad establecidos en las normas estandarizadas. En este sentido, con las herramientas del aprendizaje no supervisado se pueden descubrir falencias que permitan proponer nuevos enfoques a favor de la accesibilidad web.

CONCLUSIONES

Es bien sabido que los métodos estadísticos permiten el análisis de los datos de una manera muy eficiente y de la misma manera se puede observar en el análisis de las relaciones que tienen las variables de estudio. Es importante señalar que el investigador tiene la libertad de elegir las técnicas y métodos estadísticos que crea conveniente, pero esta elección es uno de los principales puntos que marcan la calidad de los resultados, es por eso, que conocer las ventajas, desventajas y escenarios de la aplicación de las técnicas es muy relevante en un proyecto de investigación.

Las técnicas o métodos de machine learning tienen una gran importancia en la aplicación cuando en la investigación hay que evaluar un gran número de variables. Las técnicas de clasificación, agrupamiento y reducción de dimensiones son las protagonistas si se aplican para el tratado de los datos y para la obtención de respuestas en lo que es accesibilidad web. Sin embargo, se debe tener presente que cuando se habla de la web tenemos que tener claro cuál es el tipo de contenido que se va a evaluar y dependiendo de eso aplicamos los métodos.

Aplicar las técnicas de agrupamiento como el PCA, k-means, agrupación jerárquica, las técnicas de clasificación o de reducción de dimensiones, facilitan el proceso de interpretar y mostrar los datos en gráficos minimizando la incertidumbre para determinar conclusiones. En este sentido, cuando se estudia la accesibilidad web se maneja un conjunto amplio de variables que al utilizar cualquier técnica o métodos de machine learning, los resultados se convierten en parte primordial teniendo incidencia en la efectividad del análisis.

Los modelos de aprendizajes automáticos basados en las técnicas de machine learning son usados por los investigadores después de tratar los datos, por lo general con alguna técnica de análisis factorial o de reducción de dimensiones que forman parte del aprendizaje no supervisado. Luego de ese procedimiento se crea un modelo para lograr predecir los niveles de accesibilidad que tienen los portales web y de esta manera estos métodos permiten que al estudiar la accesibilidad sea mucho más fácil comprender los lineamientos de las normas aplicadas al contenido web.

Es bien sabido que los métodos estadísticos permiten el análisis de los datos de una manera muy eficiente y de la misma manera se puede observar en el análisis de las relaciones que tienen las variables de estudio. Es importante señalar que el investigador tiene la libertad de elegir las técnicas y métodos estadísticos que crea conveniente, pero esta elección es uno de los principales puntos que marcan la calidad de los resultados, es por eso, que conocer las ventajas, desventajas y escenarios de la aplicación de las técnicas es muy relevante en un proyecto de investigación.

Las técnicas o métodos de machine learning tienen una gran importancia en la aplicación cuando en la investigación hay que evaluar un gran número de variables. Las técnicas de clasificación, agrupamiento y reducción de dimensiones son las protagonistas si se aplican para el tratado de los datos y para la obtención de respuestas en lo que es accesibilidad web. Sin embargo, se debe tener presente que cuando se habla de la web tenemos que tener claro cuál es el tipo de contenido que se va a evaluar y dependiendo de eso aplicamos los métodos.

Aplicar las técnicas de agrupamiento como el PCA, k-means, agrupación jerárquica, las técnicas de clasificación o de reducción de dimensiones, facilitan el proceso de interpretar y mostrar los datos en gráficos minimizando la incertidumbre para determinar conclusiones. En este sentido, cuando se estudia la accesibilidad web se maneja un conjunto amplio de variables que al utilizar cualquier técnica o métodos de machine learning, los resultados se convierten en parte primordial teniendo incidencia en la efectividad del análisis.

Los modelos de aprendizajes automáticos basados en las técnicas de machine learning son usados por los investigadores después de tratar los datos, por lo general con alguna técnica de análisis factorial o de reducción de dimensiones que forman parte del aprendizaje no supervisado. Luego de ese procedimiento se crea un modelo para lograr predecir los niveles de accesibilidad que tienen los portales web y de esta manera estos métodos permiten que al estudiar la accesibilidad sea mucho más fácil comprender los lineamientos de las normas aplicadas al contenido web.

REFERENCIAS

- [1] M. Campoverde-Molina, S. Luján-Mora, and L. Valverde, "Análisis de accesibilidad web de las universidades y escuelas politécnicas del Ecuador aplicando la norma NTE INEN ISO/IEC 40500:2012," pp. 53-68, 2019.
- [2] A. Ismail and K. S. Kuppusamy, "Web accessibility investigation and identification of major issues of higher education websites with statistical measures: A case study of college websites," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 3, pp. 901-911, 2022.
doi: 10.1016/j.jksuci.2019.03.011.
- [3] The World Wide Web Consortium, "Home | Web Accessibility Initiative (WAI) | W3C." <https://www.w3.org/WAI/> (accessed jul. 31, 2022).
- [4] D. Naranjo-Villota, J. Guña-Moya, P. Acosta-Vargas, and V. Muirragui-Irrazábal, "Evaluación de la accesibilidad web en institutos acreditados de educación superior del Ecuador," Rev. Espac., vol. 41, no. 4, p. 5, 2020, [Online]. Available: <http://revistaespacios.com/a20v41n04/20410405.html>
- [5] T. Acosta, P. Acosta-Vargas, J. Zambrano-Miranda, and S. Lujan-Mora, "Web Accessibility Evaluation of Videos Published on YouTube by Worldwide Top-Ranking Universities," IEEE Access, vol. 8, pp. 110994-111011, 2020, doi: 10.1109/ACCESS.2020.3002175.

- [6] N. K. Esparza Cruz, Z. Merino Acosta, and H. Guerrero Torres, "Accesibilidad Web en las Instituciones de Educación Superior del Ecuador: Año 2016," *J. Sci. Res. Rev. Cienc. e Investig.*, vol. 1, no. CITT2016, pp. 44–48, 2016, doi: 10.26910/issn.2528-8083vol1isscitt2016.2016pp44-48.
- [7] Ó. R. G. López, T. M. B. Palacios, and M. B. Mateos, "El índice cuantitativo de calidad web como instrumento objetivo de medición de la calidad de sitios web corporativos," *Investig. Eur. Dir. y Econ. la Empresa.*, vol. 19, no. 1, pp. 16–30, 2013, doi: 10.1016/j.iedee.2012.07.004.
- [8] Y. S. Rodríguez, E. Á. Calderón, L. B. Pérez, and C. A. S. Anlas, "State of web accessibility of e-government portals in Latin America," *Bibl. An. Investig.*, vol. 16, no. 1, pp. 7–22, 2020.
- [9] L. Olsina and M. Bérnabe, "Técnicas Estadísticas para el Análisis de la Calidad de Sitios Web," no. July, 2019.
- [10] D. Sato, H. Takagi, and C. Asakawa, "Accessibility evaluation based on machine learning technique," *Eighth Int. ACM SIGACCESS Conf. Comput. Access. ASSETS 2006*, vol. 2006, pp. 253–254, 2006, doi: 10.1145/1168987.1169041.
- [11] M. Zhang, C. Wang, Z. Yu, C. Shen, and J. Bu, "Active learning for web accessibility evaluation," *Proc. 14th Web All Conf. W4A 2017*, 2017, doi: 10.1145/3058555.3058559.
- [12] S. Bahram, D. Sen, and R. S. Amant, "Prediction of Web page accessibility based on structural and textual features," *W4A 2011 - Int. Cross-Disciplinary Conf. Web Access.*, pp. 5–8, 2011, doi: 10.1145/1969289.1969329.
- [13] I. Hernández, C. R. Rivero, D. Ruiz, and R. Corchuelo, "A statistical approach to URL-based web page clustering," *WWW'12 - Proc. 21st Annu. Conf. World Wide Web Companion*, pp. 525–526, 2012, doi: 10.1145/2187980.2188109.
- [14] M. Usama et al., "Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges," *IEEE Access*, vol. 7, pp. 65579–65615, 2019, doi: 10.1109/ACCESS.2019.2916648.
- [15] J. P. Bigham, "Increasing web accessibility by automatically judging alternative text quality," *Int. Conf. Intell. User Interfaces, Proc. IUI*, pp. 349–352, 2007, doi: 10.1145/1216295.1216364.
- [16] P. Taylor, C. Fraley, and A. E. Raftery, "Journal of the American Statistical Association and Density Estimation Model-Based Clustering, Discriminant Analysis, and Density Estimation," no. May, pp. 37–41, 2012.
- [17] T. Gowda and C. Mattmann, "Clustering web pages based on structure and style similarity," *Proc. - 2016 IEEE 17th Int. Conf. Inf. Reuse Integr. IRI 2016*, pp. 175–180, 2016, doi: 10.1109/IRI.2016.30.
- [18] J. Mucha, M. Snaprud, and A. Nietzio, "Web page clustering for more efficient website accessibility evaluations," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9758, pp. 259–266, 2016, doi: 10.1007/978-3-319-41264-1_35.
- [19] M. D. Olvera Lobo, M. Aguilar Soto, and E. Ruíz de Osma, "Evaluación de sitios web de postgrados biomédicos en España Evaluation of websites for biomedical," vol. 24, no. 1, pp. 47–60, 2012.
- [20] Servicio Ecuatoriano de Normalización (INEN), "Reglamento Técnico Ecuatoriano RTE INEN 288 'Accesibilidad para el Contenido Web,'" pp. 1–6, 2016, [Online]. Available: <http://www.normalizacion.gob.ec/wp-content/uploads/downloads/2016/02/RTE-288.pdf>.