

**Uso de algoritmos Machine Learning en la clasificación de objetos astronómicos: una
revisión sistemática**

Maria Fernanda Forero Dorado

Maria Alejandra Hernández Castañeda

Asesor

Nidia Danigza Lugo López

Universidad Nacional Abierta y a Distancia – UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería – ECBTI

Especialización en Ciencia de Datos y Analítica

2024

Dedicatoria

En primer lugar, a Dios, por permitirnos vivir esta etapa de aprendizaje y desarrollo profesional.

A nuestras familias, por su apoyo constante y motivación para avanzar. Y a todas las personas que hicieron parte de este proceso.

Agradecimientos

A la directora de proyecto de grado, por su apoyo y acompañamiento constante durante el proceso. A la Universidad Nacional Abierta y a Distancia, por los conocimientos adquiridos y las herramientas que nos brindaron. A los compañeros con quienes trabajamos durante el programa, por sus conocimientos y apoyo. Y finalmente, a nuestras familias por creer en nosotras.

Resumen

La clasificación de los discos protoplanetarios se ha convertido en una tarea de vital importancia para determinar su estructura y si realmente se trata de estos cuerpos. Puesto que se está experimentando un enorme aumento de datos debido a los avances en las tecnologías telescópicas con salidas digitales automatizadas, la cantidad de datos ha aumentado exponencialmente, con una generación de aproximadamente 15 TB de datos por noche, lo que, a su vez, ha aumentado la demanda de análisis de datos complejos para una toma de decisiones adecuada. El principal objetivo de este trabajo es explorar la literatura durante los últimos diez (10) años y encontrar métodos de análisis y clasificación de discos protoplanetarios mediante diferentes algoritmos machine learning, clasificarlas y evaluar la precisión y las limitaciones estas comparado con los métodos tradicionales, a la vez que se identifican patrones en común en el uso de las distintas técnicas, dependiendo del cuerpo astronómico, el lugar de estudio y el cambio de estas técnicas con el paso de los años. Si bien existen avances prometedores en el análisis de los discos protoplanetarios, es aún un desafío incorporar el uso de los modelos de machine learning que se aplican en el estudio de otros cuerpos a los discos protoplanetarios. Sin embargo, esto no es una limitante para que se puedan modificar y adecuar los modelos ya existentes de estrellas y galaxias, por ejemplo, a los discos.

Palabras claves: Discos protoplanetarios, técnicas machine learning, exploración, tecnologías telescópicas.

Abstract

The classification of protoplanetary disks has become a task of vital importance to determine their structure and confirm their nature. Due to the rapid advancements in telescopic technologies with automated digital outputs, there has been an exponential increase in data generation, reaching approximately 15 TB of data per night. This has, in turn, heightened the demand for complex data analysis to enable proper decision-making. The main objective of this study is to explore literature from the past ten (10) years and identify methods for the analysis and classification of protoplanetary disks using various machine learning algorithms. These methods are compared and evaluated in terms of their accuracy and limitations against traditional techniques, while also identifying common patterns in the use of different techniques depending on the astronomical body, the location of the study, and how these techniques have evolved over time. While there are promising advances in the analysis of protoplanetary disks, incorporating the use of machine learning models applied to the study of other celestial bodies into the analysis of these disks remains a challenge. However, this is not a limitation, as existing models for stars and galaxies, for example, can be modified and adapted for use with protoplanetary disks.

Keywords: Protoplanetary disks, machine learning techniques, exploration, telescopic technologies.

Tabla de Contenido

Introducción	10
Planteamiento del Problema	11
Justificación	13
Objetivos	15
Objetivo General	15
Objetivos Específicos	15
Marco Conceptual	16
Física y Formación de los Discos	16
Historia del Análisis de los Discos Protoplanetarios	20
Revisión de Tendencias en el Análisis y Clasificación de Cuerpos Astronómicos	20
Análisis Bibliométrico con Vosviewer	25
Nodos Identificados	27
Análisis de los Nodos Identificados	30
Clasificación de las Diferentes Técnicas Machine Learning Usadas en el Estudio de Discos Protoplanetarios y Otros Cuerpos Celestes	34
Clasificación de las Técnicas de Redes Neuronales	37
Técnicas Basadas en Redes Neuronales	37
Técnicas de Aprendizaje Supervisado	38
Técnicas de Aprendizaje No Supervisado	39
Técnicas de Aprendizaje por Refuerzo	39
Técnicas de Procesamiento del Lenguaje Natural	39

Métodos de Ensamble	40
Técnicas Probabilísticas y de Ingeniería	40
Discusión de Resultados	40
Aplicación de Técnicas de Machine Learning en la Clasificación de Cuerpos Astronómicos y Discos Protoplanetarios	42
Redes Neuronales.....	42
Aprendizaje Supervisado	44
Conclusiones.....	46
Recomendaciones	49
Bibliografía.....	51

Lista de Tablas

Tabla 1 <i>Palabras o Términos Claves de Búsqueda</i>	21
Tabla 2 <i>Criterios y Limitaciones de la Búsqueda</i>	22
Tabla 3 <i>Tesaurus Aplicado al Análisis Bibliométrico</i>	25
Tabla 4 <i>Número de Artículos Asociados a Cada Técnica de Machine Learning</i>	35
Tabla 5 <i>Artículos Asociados a Técnicas Machine Learning por Palabras Clave</i>	36
Tabla 6 <i>Número de Artículos por Técnica de Machine Learning</i>	40

Lista de Figuras

Figura 1 <i>Proceso de Formación de una Estrella de Masa Baja</i>	17
Figura 2 <i>Etapas de Formación de un Sistema Planetario</i>	19
Figura 3 <i>Referencias por Principales Países</i>	23
Figura 4 <i>Publicaciones por Año</i>	23
Figura 5 <i>Mapa de Concurrencia de Palabras Clave con VOSViewer</i>	28

Introducción

El estudio del cosmos es algo que siempre ha inquietado a la humanidad y, con el gran avance que ha tenido la ciencia, tanto en física, química, astronomía, como en las ciencias computacionales, se han logrado conocer a mayor profundidad diferentes aspectos sobre la creación del universo, las estrellas y galaxias. Pero surge la pregunta, ¿Qué cuerpos explican el origen de las estrellas y los sistemas planetarios? La respuesta radica en los discos protoplanetarios, estructuras que desempeñan un papel crucial en la formación de estrellas y planetas.

La acelerada generación de datos gracias a los avances en tecnologías telescópicas, como las producidas por telescopios como ALMA y Hubble, ha permitido obtener imágenes más nítidas y detalladas de estos cuerpos, como lo son los detalles de las partículas de polvo en un disco protoplanetario (ALMA, 2023). No obstante, a pesar de los avances en el campo, la aplicación de algoritmos de *machine learning* en el estudio de discos protoplanetarios sigue siendo un área poco explorada en comparación con otras investigaciones astronómicas más generalizadas, como la clasificación de espectros estelares o la estimación de parámetros atmosféricos estelares. Con este trabajo se quiere proporcionar una visión de las metodologías actuales y así, analizar las oportunidades que presenta este campo para futuras investigaciones.

Este trabajo pretende identificar qué herramientas de machine learning se aplican actualmente en el análisis y clasificación de diferentes objetos astronómicos, mediante una revisión bibliográfica de artículos de investigación. Así mismo, se busca observar si las técnicas aplicadas al estudio de otros objetos astronómicos pueden aplicarse a los discos protoplanetarios.

Planteamiento del Problema

Pese a los grandes avances de la ciencia en las últimas décadas en el reconocimiento e identificación de cuerpos astronómicos y el uso cada vez más amplio de herramientas de machine learning e inteligencia artificial (Rodríguez et al, 2022), actualmente se cuenta con poca información acerca de la aplicación de estas técnicas en el reconocimiento y estudio de los discos protoplanetarios, los cuales son parte fundamental del origen de las estrellas y planetas (Forbes et al, 2019).

Dada la falta de referencias bibliográficas enfocadas en el estudio de estos cuerpos, los investigadores encuentran dificultades para iniciar o continuar sus investigaciones, por la falta de bases específicas. Parte del problema puede deberse, a que la aplicación de los algoritmos de aprendizaje automático tuvo su auge en la década pasada, impulsado por los datos recopilados por telescopios como el ALMA, Hubble, entre otros, que permiten obtener imágenes más claras de las que se obtenían previamente (Kremer et al, 2017).

Algunos científicos han documentado la aplicación de los algoritmos de manera específica para los discos protoplanetarios, pero el campo de investigación no es ampliamente estudiado, contrario a temas más generales, como la clasificación de espectro estelar, redshifts fotométricos, estimación de parámetros estelares atmosféricos, entre otros (Reza, 2021). En este orden de ideas, se tiene un campo del conocimiento aún por explorar y con amplias oportunidades de investigación y aprendizaje.

Algunas de estas aplicaciones en temáticas generales pueden ser estudiadas, evaluadas y, si resultan convenientes, aplicadas al estudio de los discos protoplanetarios, considerando que estos comparten algunas características con otros cuerpos astronómicos como la temperatura, densidad, tamaño, que permiten clasificarlos y estudiarlos a mayor profundidad. En este sentido,

el avance de herramientas computacionales, como las técnicas de machine learning, ha demostrado ser útil para el análisis de grandes volúmenes de datos en distintas áreas de la astronomía. ¿Cómo puede el uso de técnicas de machine learning optimizar la clasificación y el estudio de los discos protoplanetarios en función de estas características?

Justificación

Dentro de las aplicaciones de la ciencia de datos en la astronomía, se observa un crecimiento acelerado en los últimos años de la aplicación de metodologías de machine learning, pasando de casi cero en la década de 1990 a alrededor de 35000 a inicios de la década del 2020 (Rodríguez et al, 2022), lo que demuestra que, de la mano con el avance de la tecnología se ha dado un avance de la astronomía y astrofísica.

Contrario a este comportamiento, en la búsqueda de información relacionada con la aplicación de técnicas de machine learning (ML) a los discos protoplanetarios, el número de publicaciones es bastante bajo. Las publicaciones, artículos y libros disponibles se enfocan en analizar las características fisicoquímicas de los discos, a partir de observaciones de telescopios y mapas de calor, más no en el uso, aplicaciones o resultados de metodologías de análisis de datos per se. Sí bien la información disponible relacionada con los discos protoplanetarios representa un punto de partida para el diseño de modelos de aprendizaje automático, este es un camino aún por recorrer.

A través de esta investigación se busca conocer las aplicaciones del ML y demás algoritmos de análisis de datos en la clasificación de objetos astronómicos, establecer sus generalidades, sus características de uso, los resultados obtenidos y, partiendo de las características puntuales conocidas de los discos protoplanetarios, categorizar las aplicaciones que se pueden emplear y los beneficios de esta aplicación, considerando el tamaño de las bases de datos que se generan a partir de las observaciones (Kremer et al, 2017, p18).

Esto es relevante, considerando que la ciencia y la tecnología continúan avanzando a grandes velocidades y, desde los campos de la astronomía y astrofísica existen múltiples aplicaciones posibles, no solo para identificar los cuerpos exoplanetarios, sino también para

generar conocimiento y que este pueda ser transmitido de manera adecuada, confiable y clara tanto a quienes hacen parte de estos campos, como a quienes encuentren interés en aprender más sobre nuestro universo observable.

Mediante este trabajo se busca aportar al campo de la astronomía al explorar un área poco investigada: la aplicación de técnicas de machine learning en el estudio de discos protoplanetarios. Al establecer un marco teórico para el uso de algoritmos de aprendizaje automático en la clasificación y análisis de los discos, se sientan las bases para futuras investigaciones que puedan mejorar la comprensión de los procesos de formación planetaria.

Objetivos

Objetivo General

Exponer el uso de las diversas aplicaciones y algoritmos de Machine Learning (ML) en el análisis, clasificación y estudio de los discos protoplanetarios.

Objetivos Específicos

Identificar la literatura científica relacionada con el uso de Machine Learning y otros algoritmos en el estudio de discos protoplanetarios y otros cuerpos celestes.

Clasificar las diferentes técnicas de Machine Learning utilizadas estudio de discos protoplanetarios y otros cuerpos celestes, como clasificadores supervisados, no supervisados y semi supervisados.

Analizar las técnicas de Machine learning aplicadas en la clasificación de otros cuerpos astronómicos y su posible uso en la clasificación de discos protoplanetarios.

Marco Conceptual

Dentro del estudio del universo ha existido siempre la pregunta de cómo se creó todo lo existente, desde los cometas hasta los planetas y estrellas. Con el paso de los años y los avances que ha tenido la ciencia, sobre todo en años recientes, se han podido descubrir diferentes cuerpos celestes y su comportamiento. Dentro de los descubrimientos ya mencionados se encuentran los discos protoplanetarios, o puntos de creación de planetas y estrellas.

Física y Formación de los Discos

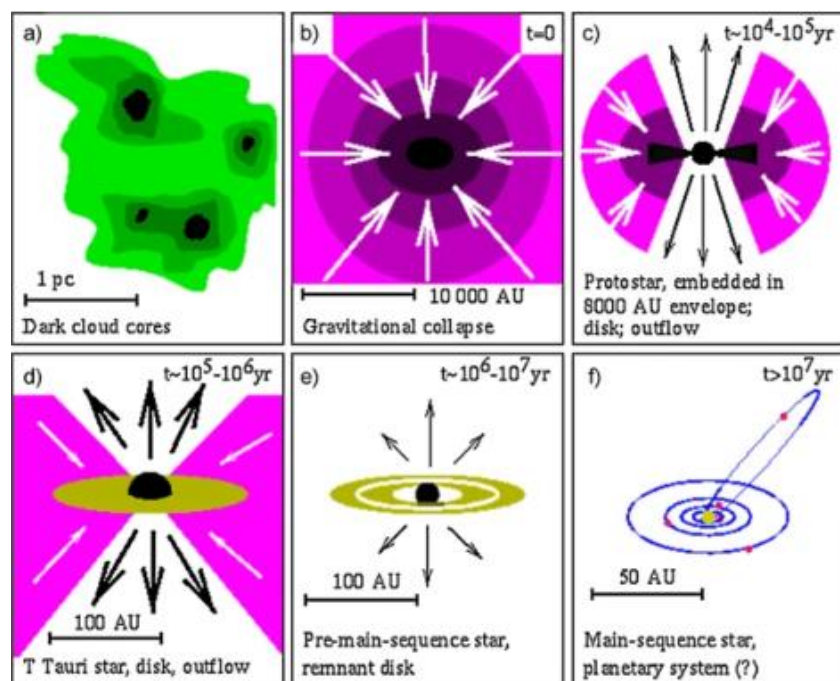
El origen de estos discos, como lo mencionan Armitage (2020) y Cánovas (2018) “se da a partir de nubes compuestas de gas (hidrógeno y helio) y partículas de polvo (formado a partir de metales) cuyo tamaño es de alrededor de una milésima parte de un milímetro” (p. 24). Estas sustancias, eventualmente llegan a formar cúmulos dentro de una nube molecular (Figura 1a) al agruparse. Estos, debido a su gravedad, llegan a colapsarse (Figura 1b) los que a su vez darán origen a la formación de estrellas, conocidas como protoestrellas (Figura 1c). En medio de este proceso se crean los discos protoplanetarios, como consecuencia del momento angular generado por la rotación de la nube alrededor de la estrella (Figura 1d). En el proceso evolutivo de la estrella, una vez esta se ha formado, queda un disco remanente a su alrededor (Figura 1e), que eventualmente desaparecerá (Figura 1f), dando paso a los sistemas planetarios alrededor de la estrella madura.

El disco se transforma a través de procesos de acreción viscosa, donde el material en el disco pierde su momento angular debido a su viscosidad microscópica, acumulándose hacia la estrella central y desplazándose hacia las partes exteriores del disco (Figura 1d). También puede presentarse la foto evaporación, que ocurre cuando la radiación llega a las capas externas del

disco y proporciona suficiente energía térmica para que el material se desprenda del campo gravitatorio de la protoestrella.

Figura 1

Proceso de Formación de una Estrella de Masa Baja



Nota. La figura muestra el proceso de formación de una estrella de masa baja. Tomado de Mandt et al., (2014).

En un disco alrededor de una estrella aislada, la radiación proviene de la estrella central, pero si la estrella no está aislada, la radiación puede venir del entorno circundante. Esto sucede porque la región de formación de los discos protoplanetarios puede incluir diferentes tipos de estrellas, incluidas estrellas masivas lo suficientemente cercanas como para que su radiación afecte al disco. Independientemente de la fuente de radiación que incide en el disco, las tasas de pérdida de masa debido a la foto evaporación se determinan resolviendo ecuaciones hidrodinámicas para un disco alrededor de una estrella (Gómez, J. 2023).

Generalmente, estos discos tienen un tamaño de 100 unidades astronómicas o au (1 au = 149.597.871 km) y presentan una rotación de tipo Kepleriana, es decir, que giran en una órbita alrededor de la estrella en su centro, como lo hacen los planetas del sistema solar entorno al sol.

Para clasificar los discos se debe considerar la pendiente de la distribución espectral de energía (DEE) o representación gráfica de la cantidad de energía emitida por la protoestrella, así como la temperatura a la que se calienta el disco, generado por la radiación recibida de la estrella. Estos factores se relacionan mediante la ecuación 1:

$$\alpha_{IR} = \frac{d \log \nu F_{\nu}}{d \log \nu} \equiv \frac{d \log \lambda F_{\lambda}}{d \log \lambda} \quad (1)$$

Donde α_{IR} representa la emisión de objetos astronómicos. La primera expresión mide como cambia el flujo con respecto a la frecuencia emitida (en escala logarítmica) y la segunda expresión es equivalente a la primera, pero para estudios donde se trabaja en términos de longitud de onda. De acuerdo con los valores obtenidos, los discos se pueden clasificar en:

Clase 0: disco poco evolucionado (Figura 2a). La distribución espectral tiene un pico en $\sim 100 \mu m$ y no se detecta flujo infrarrojo cercano. La masa que rodea la estrella es de $>0.5 M_{\odot}$ (masa solar). La edad de este objeto es de $\sim 10^4$ años.

Clase I: La forma de la distribución espectral de energía se achata y la cantidad de polvo se reduce, dejando así una masa de $\sim 0.1 M_{\odot}$. Esto indica que se tiene una estructura discal joven ($\sim 10^5$ años), y la estrella se mantiene en el centro de la nube (Figura 2b).

Clase II: Posee una distribución espectral en descenso (Figura 2c). El flujo de radiación comienza a reducirse y la estrella ha perdido la envoltura, aunque aún retiene un disco a su

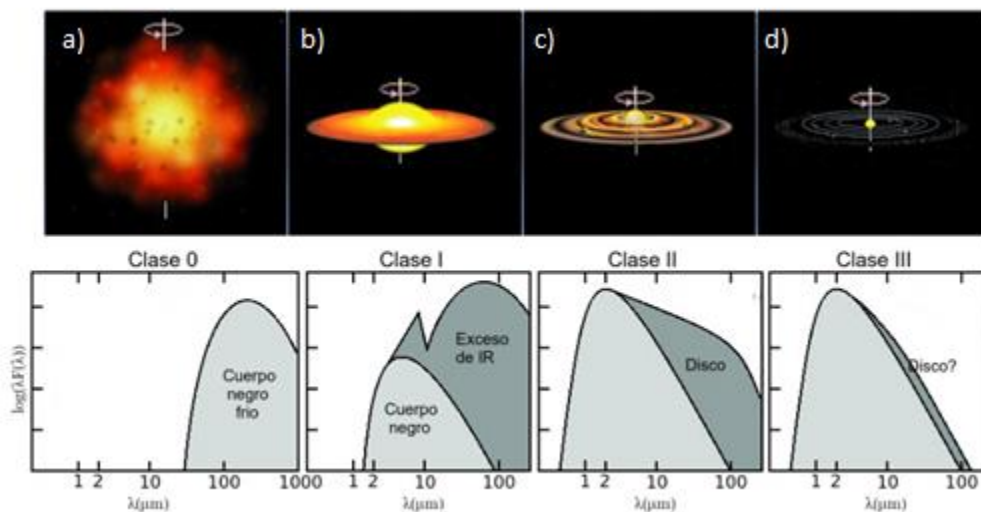
alrededor. Los objetos en esta etapa se conocen como Estrellas T-Tauri Clásicas. La masa es de $\sim 0.01 M_{\odot}$, y su edad de $\sim 10^6$ años.

Clase III: No cuentan con exceso en el flujo de radiación. La nube que rodeaba a la estrella ha desaparecido en su totalidad, por lo que su masa es ahora de $\sim 0.03 M_{\odot}$ (Figura 2d). En este punto, la edad del objeto es de $\sim 10^7$ años.

Esta clasificación se puede observar en la figura 2:

Figura 2

Etapas de Formación de un Sistema Planetario



Nota. Evolución de los sistemas estelares jóvenes, desde el colapso inicial de la nube hasta el sistema planetario maduro, junto con sus correspondientes espectros de emisión infrarroja. Sus clases reflejan diferentes etapas en la formación estelar y la dispersión de material circundante.

Tomado de *Repositorio UPTC*, Buitrago, N. 2018, p7.

La clasificación anterior corresponde a un sistema que evoluciona conforme pasa el tiempo de manera progresiva, por ello, se dificulta establecer con precisión la edad de un disco y

se prefiere dar estimaciones relativas, aunque se infiere a partir de observaciones y estudios del ALMA que la formación de anillos y surcos presentes en los discos de clase 2 empiezan a formarse cuando los discos tienen cerca de 300.000 años (Orlando et al., 2024).

Historia del Análisis de los Discos Protoplanetarios

El análisis de estos cuerpos inició alrededor de 1984 (Smith and Terrile), brindando una fuerte evidencia de la relación entre los discos circunestelares con estrellas jóvenes de tipo solar anteriores a su secuencia principal, es decir, previo a que se realice la fusión de átomos de hidrógeno para crear helio. Como menciona Strom et al (p. 1451), mediante el uso de imágenes directas y longitud de onda de infrarrojo del Satélite Astronómico Infrarrojo (IRAS), se revelaron las estructuras aplanadas de los discos asociados a 3 estrellas jóvenes: HL Tau, R Mon y L15551/IRS 5.

Posterior a esto, se demostró, mediante detectores sensibles en longitud de onda milimétrica, que muchos discos contenían granos de polvo en su estructura (Weintraub et al., 1989) con material suficiente para formar sistemas planetarios a la misma escala del sistema solar (Beckwith et al., 1990). La medición mediante interferencia de ondas de las longitudes de los discos permitió resolver las incógnitas existentes en cuanto a la rotación de estos y, la evidencia de su estructura aplanada se dio gracias al telescopio Hubble, a través de las imágenes captadas en las que se observan las sobras de discos sobre nebulosas (O'Dell and Wen, 1994).

Revisión de Tendencias en el Análisis y Clasificación de Cuerpos Astronómicos

Para la revisión bibliográfica acerca del uso de técnicas machine learning en la clasificación de objetos astronómicos se realiza la búsqueda inicial aplicando la ecuación *protoplanetary disks AND machine learning*, dado el enfoque de este trabajo. Al buscar en Web

of science se encontraron solo 22 documentos sin aplicar restricciones de años de publicación, idioma u otros.

Partiendo de estos resultados iniciales, se realizaron búsquedas y consultas en diferentes páginas y publicaciones abiertas relacionadas con la astronomía, como el observatorio ALMA, la NASA, entre otros, con el fin de tener una comprensión inicial de qué son los discos protoplanetarios y el impacto que estos tienen en la formación de los demás cuerpos astronómicos y su relación en cuanto a descubrimiento y caracterización de estos.

Por medio de esta búsqueda se identificaron los siguientes términos clave:

Tabla 1

Palabras o Términos Claves de Búsqueda

Cuerpos astronómicos	Metodologías
Protoplanetary disks (discos protoplanetarios)	Machine learning (aprendizaje automático)
Astronomical object (objetos astronómicos)	Neural networks (redes neuronales)
Exoplanets (exoplanetas)	Clasificación

Nota. Se presentan los términos clave empleados para la búsqueda de artículos.

Con estos términos se construye la ecuación de búsqueda inicial:

("Protoplanetary disks" OR "Astronomical objects") AND ("Machine learning" OR "Neural networks") AND ("Classification")

Con la ecuación ya establecida, se identifican los repositorios de información y páginas en los cuales se va a realizar la búsqueda:

- Bases de datos de información: Web of science, IEEE, Scopus, Scielo, Proquest, ScienceDirect, entre otras.

- Plataformas de investigación como Google Académico, Academia y repositorios institucionales.
- Recursos y publicaciones relacionadas con la astronomía y la astrofísica, como The Astrophysical Journal.

De manera inicial, se realiza la búsqueda con la ecuación ya definida en Web of science, obteniendo un total de 110,022 resultados. Para acotar la búsqueda a las referencias más adecuadas, se aplican los siguientes criterios y limitaciones:

Tabla 2

Criterios y Limitaciones de la Búsqueda

Criterio	Descripción
Año de publicación	Incluye los contenidos publicados entre 2020 y 2024.
Tipo de documento	Artículos. Artículos de revisión
Categorías WoS	Astronomy astrophysics. Computer Science artificial intelligence.
Idiomas	Inglés Español
Campos de investigación	Astronomy astrophysics

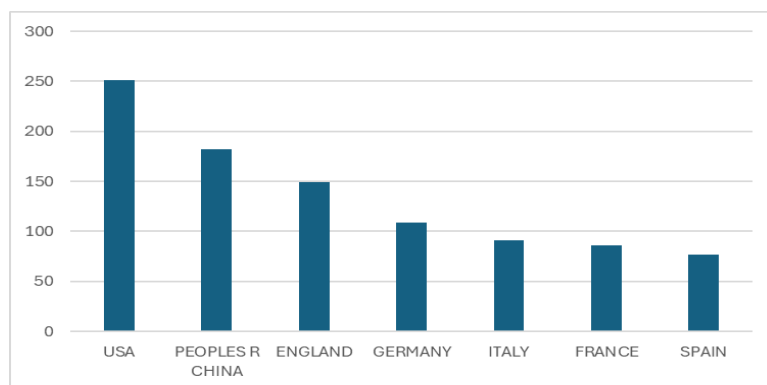
Nota. La tabla contiene el listado de limitaciones aplicada a la búsqueda realizada, para acotar los resultados.

Al aplicar estos criterios se obtuvo un total de 709 resultados. Estos resultados provienen principalmente de USA, China e Inglaterra, entre otros países que se observan en la Figura 3. Vale la pena resaltar que, en estos artículos, solo hay presencia de 3 países latinoamericanos, y Colombia no se encuentra incluida.

Así mismo, se observa en la Figura 4 el crecimiento en el número de publicaciones entre los años 2020 y 2023, siendo creciente y estando por encima de los 100 artículos por año. En 2024 se observa un comportamiento a la baja, pero vale la pena resaltar que esto puede deberse, en parte, a que al momento de realizar la consulta el año no había finalizado aún.

Figura 3

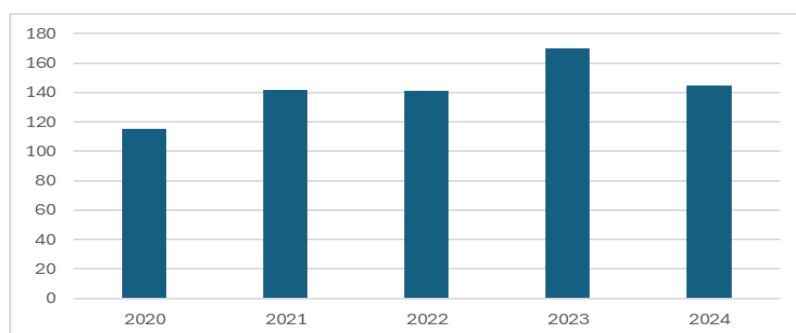
Referencias por Principales Países



Nota. Principales países que realizan publicaciones relacionadas a Machine Learning en astronomía.

Figura 4

Publicaciones por Año



Nota. Número de publicaciones realizadas por año, en el marco de tiempo definido para la revisión bibliográfica.

Adicional a esto, las áreas de interés, además de la Astrofísica astronómica son Campos de Física de partículas, geociencias multidisciplinarias, ciencias meteorológicas atmosféricas y aplicaciones interdisciplinarias de las ciencias computacionales.

Estos resultados son exportados en formato Plain Text File, para ser analizados a través del aplicativo de VosViewer. Esta información se complementa con otras publicaciones de interés y relación con el tema, que pueden consultarse a través de fuentes como Scopus y Google Académico.

Análisis Bibliométrico con Vosviewer

VOSviewer es una herramienta poderosa para el análisis y visualización de redes bibliométricas, ampliamente utilizada en investigaciones científicas. Su función es crear mapas de citas, coautoría, y co-ocurrencia de términos en bases de datos como Scopus y Web of Science, lo que permite a los investigadores obtener una percepción visual y comprensible de las relaciones entre publicaciones y autores (*VOSviewer - Visualizing Scientific Landscapes*, n.d.) (Kirby, 2023).

Al cargar la información encontrada en el aplicativo VOSviewer, esta se agrupa en clústeres dadas las palabras claves y las relaciones en las conexiones de concurrencia, formando una red de términos. Para poder crear la red o mapa, se cargan los datos extraídos en formato txt de Web of Science al aplicar la ecuación de búsqueda, sin embargo, si se crea el mapa sin realizar un procesamiento dentro de VOSviewer el resultado no será fiable, pues se incluirán términos iguales pero escritos de manera diferente, además de que resulta necesario determinar el número de concurrencias en cada palabra clave pues, si es muy bajo, muchos términos se relacionarán, aunque no deba ser de esta manera.

Para solucionar la duplicidad de términos se crea un thesaurus. Este se formula a partir de las palabras clave arrojadas por el programa, las cuales se pueden descargar como archivo .txt y editar de la siguiente manera:

Tabla 3

Tesaurus Aplicado al Análisis Bibliométrico

Término	Reemplazado por
catalog	catalogs
catalogs	catalogs
catalogues	catalogs

convolutional neural networks	convolutional neural networks
convolucional neural-networks	convolucional neural networks
galaxies	galaxies
galaxies: active	galaxies
galaxies: evolution	galaxies
galaxies: general	galaxies
galaxies: interactions	galaxies
galaxies: structure	galaxies
galaxy	galaxies
methods	methods
methods: data analysis	methods
methods: numerical	methods
methods: statistical	methods
model	model
models	model
morphological classification	morphology
morphology	morphology
neural networks	neural networks
neural-network	neural networks
neural-networks	neural networks
star-formation	stars
stars	stars
techniques	techniques
techniques: image processing	techniques
techniques: photometric	techniques
techniques: spectroscopic	techniques
variable-stars	stars
zoo	galaxies

Nota. Relación de términos existentes en la revisión bibliométrica y su reemplazo, para mejorar el análisis.

Para lograr que el tesoro funcione de manera adecuada es necesario determinar el número de ocurrencias para las palabras clave, en este caso al ver que la cantidad de datos no es significativamente grande se elige una concurrencia de 15, es decir, que aparezcan como mínimo 15 veces en las referencias utilizadas, para así asegurar que existan relaciones entre las palabras clave y que sean realmente significativas para su análisis. Finalmente se obtiene el siguiente mapa:

En total, como se observa en la Figura 4, se obtienen 5 nodos, clasificados de la siguiente manera:

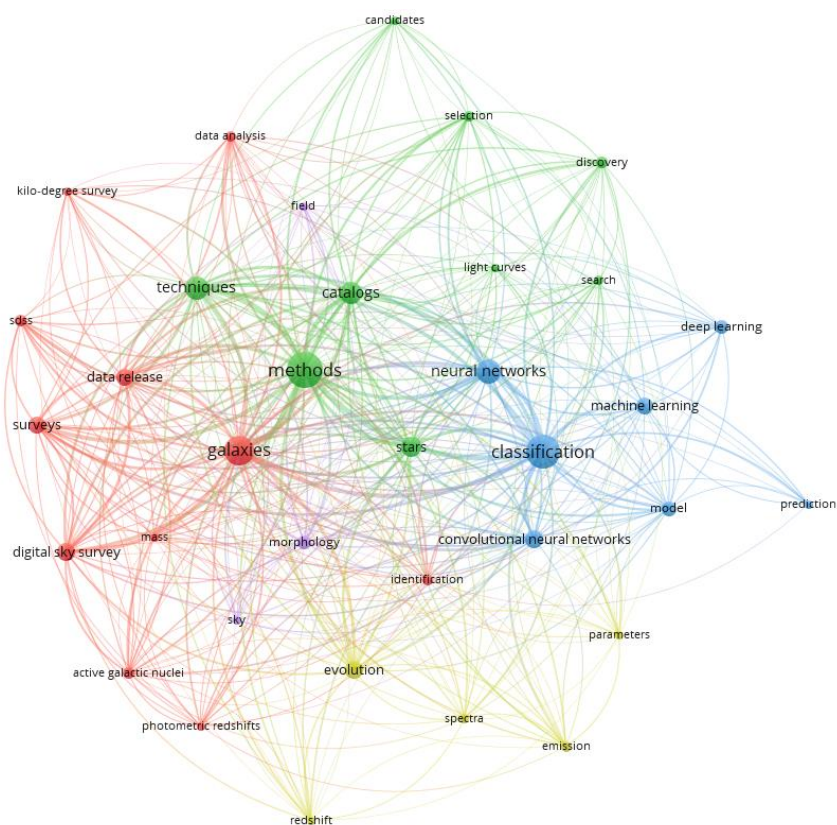
Nodos Identificados

Nodo Rojo: Su origen se encuentra en el término “galaxies”, y sus elementos son: Data “analysis”, “data release”, “digital sky survey”, “identification”, “kilo-degree survey”, “mass”, “photometric redshifts”, “sdss” y “surveys”. Los surveys son observaciones astronómicas que recopilan una amplia gama de información de los objetos celestes, y se emplean de manera general para mapear y catalogar el universo en diferentes longitudes de onda. Esto, al permitir realizar una clasificación e identificación de los objetos astronómicos, permite también la creación de catálogos de estrellas, galaxias, nebulosas, entre otros (Xu & Zhu, 2024).

Adicional a esto, los surveys han ayudado a explorar fenómenos relacionados con la materia oscura (Martins, Rosenfeld & Sobreira, 2018) y energía oscura (Wang & Zhao, 2021). Por otro lado, estudios recientes han demostrado que los datos del Kilo-degree survey en combinación con el SDSS han sido empleados para estudiar las estructuras a gran escala del universo, mediante técnicas como el análisis de correlaciones angulares y desplazamientos al redshifting (Roy et al, 2019).

Figura 5

Mapa de concurrencia de palabras clave con VOSViewer



Nota. La imagen muestra los clústeres de información obtenidos del análisis bibliométrico realizado con la herramienta VosViewer.

Nodo Verde: Con origen en el término “methods”, este nodo incluye los siguientes elementos: “candidates”, “catalogs”, “Discovery”, “light curves”, “search”, “selection”, “stars” y “techniques”. Este nodo muestra que existen métodos que permiten, a partir de búsquedas y mediante la aplicación de técnicas como el análisis de datos, descubrir nuevos objetos astronómicos, y realizar la clasificación de los candidatos evaluados. Adicionalmente, el término “Light curves”, como lo indica AAVSO (s.f.), son gráficos simples de brillo vs tiempo. Este gráfico muestra donde inicia el brillo de una estrella y su evolución a lo largo del tiempo. Con

esto, se puede realizar la clasificación de los objetos, teniendo en cuenta su edad, y la cercanía con otros cuerpos astronómicos.

Para el análisis de estas light curves se emplean métodos de parametrización para resolver problemas de degeneración de sistemas planetarios y binarios que, complementados con métodos tradicionales, permiten mejorar la identificación de planetas extrasolares (Tang et al, 2022). Estas técnicas modernas, incluyendo los algoritmos de machine learning optimizan la clasificación estelar u la detección de eventos como microlentes estelares. (Vaia, 2023).

Nodo Azul: Con su origen ubicado en el término “classification”, este nodo incluye los elementos “convolutional neural networks”, “Deep learning”, “machine learning”, “model”, “neural networks” y “prediction”. Este nodo en particular muestra el uso de técnicas de inteligencia artificial empleadas en la clasificación de objetos astronómicos. Como menciona Wang et al (2018), “las grandes cantidades de información producidas por los estudios fotográficos del cielo requieren de técnicas de reducción automatizadas para extraer información en tiempos razonables” (p. 581). Es por ello por lo que los métodos de inteligencia artificial resultan tan necesarios en el análisis de información astronómica.

Como casos de aplicación de estas técnicas con IA encontramos: sistemas que utilizan redes neuronales convolucionales para clasificar rápidamente supernovas y sus galaxias con alta precisión (Muthukrishna *et al*, 2019), el uso de datos de curvas de luz procesados en mapas de calor para entrenar redes neuronales que clasifican tipos específicos de supernovas o SCONE (Qu et al, 2021), y modelos de aprendizaje profundo que pueden diferenciar entre galaxias y estrellas, destacándose frente a métodos tradicionales (Star Galaxy Classification using Deep Learning, 2024).

Nodo Amarillo: Su origen se encuentra en el término “evolution”, e incluye los términos “emisión”, “parameters”, “redshift” y “spectra”. Este nodo se centra en el estudio de la evolución de los objetos astronómicos, mediante la consideración de parámetros fisicoquímicos y su expansión en el universo. Por ejemplo, el enfoque Line-Intensity Mapping permite mapear galaxias y el gas intergaláctico, ofreciendo información valiosa sobre la evolución química en el universo temprano (Bernal & Kovetz, 2022).

Así mismo, estudios recientes han empleado la espectroscopia avanzada para medir la distribución de elementos clave como el oxígeno, en galaxias cercanas y distantes, lo que permite entender su evolución química y como los procesos de formación estelar y explosiones de supernovas moldearon las galaxias desde sus primeras etapas (Universidad de Herfordshire, s.f.)

Nodo Morado: Su origen se encuentra en el término “morphology”, y está acompañado por los elementos “field” y “sky”. Este último nodo se relaciona de manera estrecha con el anterior, dado que la morfología de los objetos astronómicos resulta clave en el análisis de su proceso evolutivo. Adicional a esto, es importante considerar también los efectos gravitacionales que influyen en la estructura de las galaxias, como la formación de colas y puentes (Buta, 2011).

Además, la sinergia entre los enfoques observacionales y teóricos es crucial para interpretar estos datos, ya que las simulaciones numéricas de interacciones galácticas proporcionan modelos detallados que ayudan a entender cómo las galaxias evolucionan y cómo las fuerzas gravitacionales afectan su morfología (Pawlik, 2014).

Análisis de los Nodos Identificados

Del análisis realizado se deduce que el mayor enfoque de los estudios consultados se centra en la clasificación de los objetos astronómicos, principalmente mediante la aplicación de herramientas de machine learning y análisis de datos. Cabe resaltar que el término “discos

protoplanetarios” no se encontró en ninguno de los clústeres, lo que indica que la clasificación de estos mediante la aplicación de métodos de inteligencia artificial y machine learning no se encuentra lo suficientemente desarrollada a hoy, en comparación con las galaxias, agujeros negros, materia oscura, entre otros cuerpos.

Pese a lo anterior, es importante resaltar que los términos que se encuentran en el clúster amarillo, relacionados con la evolución de los cuerpos astronómicos, son también aspectos clave en el estudio de los discos protoplanetarios, teniendo en cuenta que los discos también evolucionan con el paso del tiempo y presentan diferentes características respecto a la emisión de radiación, su gravitación, entre otros (Takahashi and Muto, 2018).

Así mismo, las galaxias y estrellas, que se analizan mediante las técnicas encontradas en el clúster rojo, tienen su origen en los discos protoplanetarios, por lo que aplicar estas técnicas para su análisis resulta en encontrar también datos e información de los discos, como lo indica el modelo analítico de Khajenabi et al (2017). Por otro lado, las técnicas de clasificación que se encuentran en el clúster azul, que corresponden a técnicas de machine learning que, en la actualidad, no solo se emplean en el estudio de cuerpos como galaxias y estrellas, sino también para analizar componentes específicos de estos, como anillos (Vanderplas *et al.*, 2012; Lecun, Bengio & Hinton, 2015; AstroML, 2019), flujos de masa y energía (Ravanbakshs *et al.*, 2016; Schmelzle et al, 2017; Barchi & Sautter, 2017), entre otros.

Si bien esta información no incluye de manera explícita la relación entre las técnicas y el análisis y clasificación de los discos protoplanetarios, es evidente que su uso se puede extrapolar a estos, para conocer mejor su estructura, proceso evolutivo, características fisicoquímicas, entre otros aspectos de relevancia para la ciencia, que aún no se encuentran cubiertos, o al menos no estudiados en términos de bibliografía.

Esto abre la posibilidad de realizar estudios enfocados puntualmente en la aplicación de las técnicas de aprendizaje supervisado y no supervisado para el estudio del origen de los discos y, por consiguiente, de todos los cuerpos astronómicos que se derivan de los primeros, vinculando también a los telescopios y herramientas de análisis visual, que han tenido grandes avances en los últimos años, vinculando también la inteligencia artificial para el análisis y manejo de los datos, como lo menciona el Centro de Excelencia en Astrofísica y Tecnologías Afines de Chile.

También resulta importante resaltar que el ritmo de los descubrimientos se ha acelerado debido al aumento de la sensibilidad, alcance y longitudes de onda de los observatorios y telescopios, como el Observatorio Espacial Infrarrojo (ISO) y el Telescopio Espacial Spitzer, ha permitido ampliar el inventario de discos conocidos en términos de masa estelar central, edad, entorno y estado evolutivo (Williams & Cieza, 2010, p. 3). Estos avances han incrementado el número de datos adquiridos mediante diferentes tecnologías telescópicas, haciendo muy compleja la tarea de identificar patrones ocultos detrás del conjunto de datos para los humanos individuales por lo que se han implementado técnicas de big data y métodos de aprendizaje automático para una predicción de calidad (Sen et al., 2022). Un ejemplo de ello es lo que Kaeufer et al. (2023) hacen mediante la creación de redes neuronales (NN) para emular el proceso de generación de DEE para mejorar el análisis y reducir el costo computacional.

Otro problema observado es la interacción disco-planeta, donde se quiere ajustar los modelos de espacios planetarios (simulaciones o representaciones teóricas) a los espacios observados en los discos protoplanetarios para revelar la presencia de planetas ocultos, sin embargo, realizar un ajuste personalizado para cada objetivo resulta poco práctico debido a la complejidad de esta interacción, para esto Auddy y Lin (2020) proponen el Disk Planet Neural

Network (DPNNNet), un modelo eficiente de brechas planetarias que explota el poder del aprendizaje automático para su clasificación.

Terry et al. (2022) proponen un modelo de aprendizaje automático que puede detectar de manera rápida y precisa la presencia de uno o más planetas y limitar correctamente su ubicación.

Otra de las tareas que sea ha querido mejorar mediante el uso de machine learning es la clasificación de los discos protoplanetarios, Basak et al. (2021) proponen explorar la eficacia del aprendizaje automático (ML) para caracterizar exoplanetas en diferentes mediante arquitecturas de redes neuronales para la tarea de clasificación.

Clasificación de las Diferentes Técnicas Machine Learning Usadas en el Estudio de Discos Protoplanetarios y Otros Cuerpos Celestes

Como se observó en el segmento anterior, no se encontró de manera explícita la relación entre las técnicas y el análisis y clasificación de los discos protoplanetarios, pero si se identifica la aplicación de diferentes técnicas para clasificación de los objetos astronómicos. Con el archivo generado extraído de web of science al aplicar la ecuación de búsqueda y filtros mencionados en el capítulo anterior, se hace un análisis manual de las fuentes bibliográficas, buscando los términos relacionados a machine learning, identificando así las siguientes técnicas de machine learning:

1. Neural Network.
2. Convolutional Neural Network.
3. Deep Learning.
4. Random Forest.
5. Support Vector Machine.
6. Clustering.
7. Decision Tree.
8. Transfer Learning.
9. Recurrent Neural Network.
10. Gradient Boosting.
11. Logistic Regression.
12. Dimensionality Reduction.
13. K-Nearest Neighbors.
14. Unsupervised Learning.

15. Feature Engineering.
16. Linear Regression.
17. Reinforcement Learning.
18. Natural Language Processing.
19. Ensemble Methods.
20. Bayesian Networks.

Luego de determinar las técnicas, se pretende observar cuál de ellas es la más usada en la clasificación de los objetos astronómicos. Para ello se crea un código en Python que permite identificar cuántos artículos están asociados a cada técnica de Machine Learning, teniendo en cuenta las siguientes variables: article title (título del artículo), author keywords (palabras clave de los autores), keywords plus (palabras clave adicionales), abstract (resumen), obteniendo el siguiente resultado:

Tabla 4

Número de Artículos Asociados a Cada Técnica de Machine Learning

Técnica	Conteo de Artículos
Neural Network	566
Convolutional neural network	302
Deep learning	214
Random Forest	67
Support vector machine	46
Clustering	37
Decision Tree	24
Transfer learning	23
Supervised learning	20
Recurrent neural network	19
Gradient Boosting	16

Logistic regression	15
Dimensionality reduction	15
K-nearest neighbors	7
Unsupervised learning	6
Feature engineering	5
Linear regression	4
Reinforcement learning	3
Natural language processing	3
Ensemble methods	2
Bayesian networks	1

Nota. La tabla presenta la relación del número de artículos asociados a cada técnica de Machine Learning.

La suma de menciones o número de relaciones con cada técnica es de 1395, superando el valor de artículos analizados, esto, debido a que muchos de estos artículos aplican o mencionan en alguna parte de las variables seleccionadas para el análisis una o más técnicas de machine learning. Para tener una mejor visualización del número de artículos que aplican determinada técnica, se realiza el análisis ahora teniendo en cuenta únicamente las variables author keywords (palabras clave de los autores) y keywords plus (palabras clave adicionales), tal cual como se realizó el mapa en VosViewer, obteniendo el siguiente resultado:

Tabla 5

Artículos Asociados a Técnicas Machine Learning por Palabras Clave

Técnica	Conteo de Artículos
Neural Network	89
Convolutional neural network	48
Deep learning	40
Support vector machine	11

Random Forest	7
Decision Tree	4
Supervised learning	3
Dimensionality reduction	2
Clustering	2
Bayesian networks	1
Linear regression	1
Recurrent neural network	1
Reinforcement learning	1
Unsupervised learning	1
Gradient Boosting	1
K-nearest neighbors	1
Feature engineering	1

Nota. Relación del número de artículos asociados a cada técnica de Machine Learning, a partir de sus palabras clave.

Clasificación de las Técnicas de Redes Neuronales

Se observa que las técnicas más usadas para la clasificación de los objetos astronómicos son las redes neuronales. Es posible clasificar estas técnicas de la siguiente manera:

Técnicas Basadas en Redes Neuronales

Neural Network (NN): Las redes neuronales son modelos computacionales inspirados en el cerebro humano, compuestas por capas de nodos o "neuronas" que procesan la información. Se utilizan en una variedad de tareas, como reconocimiento de voz, imágenes y más. Dentro de sus aplicaciones se encuentran la evaluación de arquitecturas y efectividad en datasets grandes y complejos (Kyselica et al, 2022), estudios que exploran la eficacia de modelos pre entrenados de CNN para clasificar objetos astronómicos en el SDSS (Jahed & Sarvari, 2024) y la clasificación

de quasars, galaxias y estrellas, mediante un enfoque multimodal que combina datos en formato imagen y metadatos tabulados para mejorar la clasificación de objetos (Ethiraj & Bolla, 2022).

Convolutional Neural Network (CNN): Las CNN son un tipo de red neuronal especializada en el procesamiento de datos con una estructura de cuadrícula, como imágenes. Se utilizan ampliamente en visión computacional (LeCun, Y., Bengio, Y., & Hinton, G, 2015).

Deep Learning: El aprendizaje profundo hace referencia a redes neuronales con muchas capas (profundas). Estas técnicas se han usado para lograr avances significativos en tareas como el reconocimiento de imágenes para aprender características jerárquicas de los datos brutos y el procesamiento del lenguaje natural para analizar, organizar y extraer información valiosa de grandes volúmenes de datos. (Goodfellow, I., Bengio, Y., & Courville, A, 2016).

Recurrent Neural Network (RNN): Las RNN son redes neuronales especializadas en trabajar con secuencias de datos, como texto o series temporales. Son utilizadas en traducción automática y predicción de series temporales.

Transfer Learning: El aprendizaje por transferencia se refiere a reutilizar un modelo preentrenado en una tarea diferente, lo que mejora el rendimiento cuando se dispone de pocos datos para la tarea específica (Pan, S. J., & Yang, Q., 2010).

Técnicas de Aprendizaje Supervisado

Random Forest: El Random Forest es un algoritmo de ensemble que utiliza múltiples árboles de decisión para mejorar la precisión del modelo y evitar el sobreajuste.

Support Vector Machine (SVM): SVM es un método que encuentra el hiperplano que mejor separa los datos en clases. Es muy popular en clasificación y regresión.

Decision Tree: Los árboles de decisión son modelos de predicción que dividen el espacio de características en regiones homogéneas.

Gradient Boosting: Este es un método de ensemble que combina modelos débiles para crear un modelo fuerte, utilizando un enfoque secuencial de corrección de errores (Friedman, J. H., 2001).

Logistic Regression: Es un modelo lineal utilizado para la clasificación binaria, donde la salida es una probabilidad.

Linear Regression: La regresión lineal es un método utilizado para modelar la relación entre variables dependientes e independientes.

K-Nearest Neighbors (KNN): Este algoritmo clasifica los puntos de datos según los puntos más cercanos en el espacio de características.

Técnicas de Aprendizaje No Supervisado

Clustering: El clustering agrupa datos en subconjuntos de acuerdo con su similitud.

Dimensionality Reduction: La reducción de dimensionalidad busca reducir el número de variables en un conjunto de datos manteniendo la información relevante.

Unsupervised Learning: El aprendizaje no supervisado implica modelos que aprenden de datos no etiquetados para encontrar patrones en ellos.

Técnicas de Aprendizaje por Refuerzo

Reinforcement Learning: El aprendizaje por refuerzo se centra en cómo los agentes deben actuar en un entorno para maximizar las recompensas acumuladas.

Técnicas de Procesamiento del Lenguaje Natural

Natural Language Processing (NLP): El procesamiento de lenguaje natural se refiere a la interacción entre las computadoras y el lenguaje humano, especialmente la comprensión de texto (Manning, C. D., & Schütze, H., 1999).

Métodos de Ensemble

Ensemble Methods: Los métodos de ensemble combinan varios modelos para mejorar el rendimiento. Ejemplos incluyen Random Forest y Gradient Boosting.

Técnicas Probabilísticas y de Ingeniería

Bayesian Networks: Las redes bayesianas son modelos gráficos que representan dependencias.

Discusión de Resultados

Después de clasificar las técnicas se construye la siguiente tabla:

Tabla 6

Número de Artículos por Técnica de Machine Learning

Técnica	# De artículos que la usan
Técnicas Basadas en Redes Neuronales	184
Técnicas de Aprendizaje Supervisado	28
Técnicas de Aprendizaje No Supervisado	5
Técnicas de Aprendizaje por Refuerzo	1
Técnicas de Procesamiento del Lenguaje Natural	0
Métodos de Ensemble	0
Técnicas Probabilísticas y de Ingeniería	2

Nota. Se presenta la cantidad de artículos por cada grupo de técnicas de Machine Learning

Las técnicas más usadas para el análisis y clasificación de los objetos astronómicos son las que están basadas en redes neuronales, las cuales se inspiran en la estructura y funcionamiento del cerebro humano, permitiendo que las máquinas aprendan patrones y características complejas a partir de grandes cantidades de datos lo que las hace ideales para la automatización y mejora de la clasificación de objetos astronómicos, lo que facilita el análisis de

grandes cantidades de datos provenientes de observatorios y telescopios. Su capacidad para identificar patrones en imágenes, espectros y señales ha permitido una clasificación más rápida y precisa, impulsando los avances en la comprensión del universo. Con el desarrollo continuo de arquitecturas más complejas como las CNNs y RNNs, las redes neuronales seguirán siendo una herramienta crucial en la astronomía moderna (ResearchGate Study on Disk Evolution, 2023).

También se observa que el aprendizaje supervisado es ampliamente aplicado en el análisis y clasificación de objetos astronómicos. Aquí, un modelo es entrenado utilizando un conjunto de datos etiquetado, donde las entradas (como características de los objetos astronómicos) están asociadas con etiquetas o categorías específicas. Este enfoque se utiliza para entrenar modelos que puedan predecir o clasificar nuevos objetos basándose en las características aprendidas a partir de datos previos lo que puede ayudar a identificar los discos protoplanetarios en un futuro al poder seleccionar características únicas de estos cuerpos y predecir o identificar su aparición en grandes cantidades de datos.

Aplicación de Técnicas de Machine Learning en la Clasificación de Cuerpos Astronómicos y Discos Protoplanetarios

En el capítulo anterior se clasificaron las técnicas Machine Learning y se identificaron las más usadas para la clasificación de diferentes cuerpos astronómicos, estas técnicas son: *redes neuronales y aprendizaje supervisado*.

Redes Neuronales

El uso de redes neuronales en la clasificación de objetos astronómicos permite aprender patrones complejos en grandes volúmenes de datos, lo que es esencial en un campo como la astronomía, donde los datos son vastos y multifacéticos, permitiendo la clasificación de imágenes astronómicas mediante las redes neuronales convolucionales (CNNs), las cuales facilitan la clasificación de imágenes de telescopios (Fluke & Jacobs, 2020; Zhang & Zhao, 2015; Dieleman, Willet & Dambre, 2015).

Esta técnica es especialmente útil en la clasificación de imágenes de discos protoplanetarios, pues al ser obtenidas mediante observaciones en longitudes de onda submilimétricas o infrarrojas, que revelan estructuras como los anillos, huecos y otras características del disco, pueden analizar imágenes complejas y aprender a identificar patrones que corresponden a diferentes tipos de discos protoplanetarios, basándose en características visuales como la forma, el tamaño y la presencia de huecos o anillos. Un ejemplo puede ser la clasificación de los discos protoplanetarios según su forma (disco de borde afilado, disco en espiral, disco asimétrico) o la presencia de estructuras específicas como huecos (lo que podría indicar la formación de planetas), a partir de imágenes obtenidas por telescopios espaciales como el ALMA (Barrow et al., 2023; Sloan Digital Sky Survey, 2023), estas variables físicas y químicas clave, identificadas por estos telescopios no han sido comúnmente utilizadas como

características de entrada en modelos de aprendizaje supervisado por lo que al tener estudios que identifican estas variables, pueden ser la base para lograr que los algoritmos identifiquen y clasifiquen estos cuerpos, algunas de las más relevantes incluyen:

Densidad del gas: La densidad del gas en distintas regiones del disco es un parámetro crucial. Se ha demostrado que tiene una correlación significativa con las abundancias de moléculas como el CO, lo cual es vital para comprender la química dentro del disco (Trierweiler, 2024; Smirnov-Pinchukov et al., 2020).

Temperatura: La temperatura en el disco, tanto del gas como del polvo, influye en las reacciones químicas y en los procesos de formación planetaria. Los perfiles de temperatura varían radialmente y con la altura sobre el plano medio del disco, siendo un factor esencial para modelar su evolución (Mao et al., 2023; Benítez-Llambay & Masset, 2016).

Composición química: Las abundancias de compuestos específicos, como el CO, H₂O y otras moléculas relevantes, son indicadores clave en la dinámica química de los discos protoplanetarios. Las variaciones en la química dependen de las condiciones locales de temperatura y densidad (Padovani et al., 2018; Aikawa et al., 2021).

Radiación UV y rayos X: La radiación de la estrella central, tanto en el rango UV como en rayos X, juega un papel determinante en la evolución del disco, afectando tanto la temperatura como las reacciones químicas en su interior (Bruderer et al., 2009; Trierweiler, 2024).

Distribución radial y vertical: La geometría del disco, definida por su distribución radial y la estructura vertical, influye en la dinámica del material y en los procesos físicos y químicos que ocurren en su interior (Mao et al., 2023; Benítez-Llambay & Masset, 2016).

Aprendizaje Supervisado

El aprendizaje supervisado es ampliamente utilizado para la clasificación de objetos astronómicos en distintas categorías, como galaxias (Sabeesh & Bolla, 2020), estrellas (Karachentsev et al, 2017), nebulosas (Brunt et al, 2021), quásares (Toba et al, 2019) y exoplanetas (Ohta et al, 2018). Los modelos supervisados aprenden de un conjunto de datos etiquetado (por ejemplo, imágenes o espectros) para identificar características clave de cada tipo de objeto. A partir de datos observacionales, los modelos supervisados pueden ser entrenados para identificar discos protoplanetarios según sus características físicas como la masa del disco, su tamaño, la distribución del material y la temperatura.

Los modelos de aprendizaje supervisado como árboles de decisión, máquinas de soporte vectorial (SVM) y redes neuronales profundas pueden clasificar discos basándose en estas características y variables clave mencionadas anteriormente, por ejemplo, un modelo supervisado podría clasificar discos en función de su densidad, tamaño y forma, ayudando a predecir cuáles tienen más probabilidades de evolucionar hacia un sistema planetario estable (Morales et al., 2021); (Disk Detective Collaboration, 2022).

La aplicación de un modelo de árbol de decisión se basa en la división iterativa de características en regiones más simples según reglas lógicas, las cuales permiten construir modelos interpretables que asignan etiquetas de clasificación a cada disco. Para datos de discos protoplanetarios, un árbol podría usar umbrales de temperatura y densidad para distinguir entre discos ricos en moléculas como CO o H₂O (Smirnov-Pinchukov et al., 2020). Por otro lado, las SVM identifican un hiperplano óptimo que separa las clases dentro del espacio de características, pudiendo separar discos dominados por diferentes compuestos químicos (como CO versus CH₄) basándose en características como la radiación UV o la densidad del gas. La flexibilidad de las

SVM radica en su capacidad para usar funciones kernel, lo que les permite manejar relaciones no lineales en los datos (Mao et al., 2023). Otro ejemplo de aplicación de un aprendizaje supervisado son las redes neuronales profundas (DNN), estas son especialmente útiles para detectar patrones no lineales complejos en grandes conjuntos de datos.

En el caso de discos protoplanetarios, podrían modelar interacciones complejas entre múltiples características, como la combinación de temperatura, radiación y composición química. Por ejemplo, una DNN podría predecir si un disco tiene condiciones favorables para la formación de planetas terrestres o gigantes basándose en estas variables (Benítez-Llambay & Masset, 2016).

Conclusiones

Aunque el estudio y clasificación de los discos protoplanetarios con el uso de técnicas de machine learning no es ampliamente abarcado en la literatura científica, se observa que existen estudios enfocados en el análisis de otros cuerpos celestes que, aplicando las modificaciones necesarias, pueden llegar a extrapolarse a los discos, dado que los cuerpos cuentan con características físicas y químicas en común.

Con esta información a nivel general, se puede, de manera inicial, plantear la realización de análisis diversos sobre los discos, aplicando las técnicas de ML encontradas, como el Deep learning, neural networks y modelos predictivos para ampliar y profundizar el conocimiento del universo y, a su vez, resolver preguntas sobre la creación de las estrellas, galaxias, planteamientos sobre el futuro del universo, entre otras posibilidades.

La información encontrada es producto del uso de tecnologías avanzadas de observación, como los telescopios ALMA y Hubble, los cuales han permitido obtener imágenes detalladas de discos protoplanetarios incrementado la comprensión sobre la formación de sistemas estelares y planetarios, esto abre las puertas a realizar un mayor estudio de los discos protoplanetarios como un campo en desarrollo.

Aunque los avances en el análisis de discos protoplanetarios son prometedores, aún existen desafíos importantes, como la calidad y cantidad de los datos disponibles, la complejidad de las estructuras de los discos, y la necesidad de desarrollar modelos más robustos y específicos, es por esto que el futuro del estudio de estos discos necesitará de una fuerte colaboración interdisciplinaria entre astrónomos, ingenieros y especialistas en machine learning para superar estos obstáculos y hacer más accesible el conocimiento sobre el origen de las galaxias.

Las redes neuronales han demostrado ser herramientas poderosas para la clasificación de objetos astronómicos, aprovechando su capacidad para aprender patrones complejos en grandes volúmenes de datos, lo que es esencial en un campo como la astronomía, donde los datos son vastos y multifacéticos.

El aprendizaje supervisado posee una gran capacidad para aprender patrones a partir de grandes conjuntos de datos etiquetados, este enfoque permite una clasificación precisa y eficiente de galaxias, estrellas, exoplanetas y eventos astronómicos transitorios, lo que lo hace una herramienta esencial en la astronomía moderna.

El uso de aprendizaje supervisado en la clasificación y análisis de discos protoplanetarios ofrece una herramienta poderosa para los astrónomos. Con técnicas como las redes neuronales convolucionales (CNNs) y las máquinas de soporte vectorial (SVM), es posible clasificar discos protoplanetarios a partir de imágenes, espectros y datos multicanal, detectando estructuras internas y prediciendo la evolución de los discos. Estas técnicas ayudan a automatizar y mejorar el proceso de identificación y clasificación de discos protoplanetarios, permitiendo un análisis más eficiente y profundo de los procesos de formación planetaria en el universo.

Luego de realizar el análisis de los resultados obtenidos en VosViewer, se logró identificar que en Web of Science no se encuentra la palabra “protoplanetary disk” como palabra clave en la ecuación de búsqueda, sin embargo, al realizar el análisis de los nodos y buscar nueva información para entender mejor como los métodos existentes de clasificación de objetos astronómicos pueden ser llevados a la clasificación de discos protoplanetarios, se encontraron algunos trabajos pertenecientes a otras bases de datos como el caso de Diop et al. (2023) o Smirnov-Pinchukov et al. (2020), trabajos publicados en revistas científicas en el año 2024 a pesar de ser aceptados en años anteriores. Así mismo se encuentra que la concurrencia

entre machine learning y discos protoplanetarios es muy baja, pero si se identifican textos que se encargan de analizar, describir o clasificar estos cuerpos mediante observación de información recopilada, como el caso de la información publicada por ALMA Observatory en (2023), indicando que ya existe información suficiente para iniciar el entrenamiento de los algoritmos y ampliar la aplicación del machine learning en la clasificación de discos protoplanetarios.

Recomendaciones

Es recomendable ampliar las fuentes utilizadas en el análisis bibliométrico más allá de Web of Science, dado que la información sobre Machine Learning aplicado a discos protoplanetarios está dispersa en diversas bases de datos. A diferencia de términos más generales como "Machine Learning", que suelen estar centralizados en repositorios reconocidos, los estudios en este campo específico se encuentran distribuidos en plataformas variadas, algunas de menor alcance o especialización. Al realizar el análisis de los nodos buscando información complementaria, empleando términos específicos, se identificaron artículos relevantes en bases de datos más pequeñas o menos conocidas, que no cuentan con el mismo nivel de reconocimiento que Web of Science. Esto evidencia que restringir el análisis bibliométrico a una sola fuente puede dar lugar a una visión incompleta del panorama investigativo. Por lo tanto, resulta esencial incluir y considerar estas otras bases de datos en el análisis para garantizar una cobertura más amplia y representativa de la literatura existente.

Para avanzar en el estudio y clasificación de los discos protoplanetarios, es importante contar con bases de datos especializadas y etiquetadas de estos discos, para facilitar el entrenamiento de modelos de machine learning, dado que la disponibilidad de datos de calidad es fundamental para garantizar un buen desempeño de los modelos de aprendizaje supervisado. Así mismo, resulta importante que se desarrollen modelos específicos que consideren la complejidad y las particularidades físicas y químicas de los discos, en un trabajo colaborativo entre los astrónomos y expertos en machine learning.

Relacionado con el punto anterior y dados los grandes avances de la tecnología, es importante que se fomente la formación de equipos multidisciplinarios y programas educativos

que incluya la astronomía, la ingeniería y el machine learning, para superar los diferentes desafíos técnicos y científicos que se pueden presentar.

Finalmente, se recomienda invertir en modelos que combinen el aprendizaje supervisado con modelos no supervisados para manejar datos no etiquetados y de estructura compleja, logrando así clasificaciones más precisas y eficientes de los discos.

Bibliografía

- Aikawa, Y., et al. (2021). *Chemical evolution in protoplanetary disks*. The Astrophysical Journal, 925(2), 80-105.
- ALMA Observatory. (2023). *ALMA revela detalles de las partículas de polvo en un disco protoplanetario*. ALMA. <https://www.almaobservatory.org/es/comunicados-de-prensa/alma-revela-detalles-de-las-particulas-de-polvo-en-un-disco-protoplanetario/>
- American Association of Variable Star Observers (AAVSO). (n.d.). *About light curves*. AAVSO. <https://www.aavso.org/about-light-curves>
- Armitage, PJ. (2020). *Protoplanetary disk structure*. In *Astrophysics of planet formation*. Cambridge Univ. https://www.cambridge.org/core/services/aop-cambridge-core/content/view/1E2FD2470FCC8159D12FB8A718562051/9781108420501c1_1-48.pdf/observations_of_planetary_systems.pdf
- AstroML: Machine Learning and Data Mining for Astronomy*. (2019). Retrieved from <http://www.astroml.org>
- Auddy, S., & Lin, M. (2020). *A Machine Learning Model to Infer Planet Masses from Gaps Observed in Protoplanetary Disks*. Astrophysical Journal/The Astrophysical Journal, 900(1), 62. <https://doi.org/10.3847/1538-4357/aba95d>
- Banerjee, P., Chattopadhyay, T., Chattopadhyay, AK. (2023). *Comparison among different Clustering and Classification Techniques: Astronomical data-dependent study*. New Astronomy, 100. <https://www.sciencedirect.com/science/article/abs/pii/S1384107622001555?via%3Dihub>

- Barchi, P.H., da Costa, F.G., & Sautter, R. (2017). *Improving galaxy morphology with machine learning*. *Astronomy & Astrophysics*, 599, A120. <https://doi.org/10.1051/0004-6361/201629722>
- Barrow, R. A., Jones, P., & Nguyen, T. (2023). *Supervised learning for astrophysical data analysis*. *Astrophysical Journal Letters*, 887(4), 1-10.
- Basak, S., Mathur, A., Theophilus, A. J., Deshpande, G., & Murthy, J. (2021). *Habitability classification of exoplanets: a machine learning insight*. *The European Physical Journal. Special Topics*, 230(10), 2221-2251. <https://doi.org/10.1140/epjs/s11734-021-00203-z>
- Bassi, S., Sharma, K., Gomekar, A. (2021). *Classification of Variable Stars Light Curves Using Long Short Term Memory Network*. *Frontiers in Astronomy and Space Sciences*, 8. <https://www.frontiersin.org/journals/astronomy-and-space-sciences/articles/10.3389/fspas.2021.718139/full>
- Beckwith, S., Sargent, A., Chini, R., Guesten, R. (1990). *A Survey for Circumstellar Disks around young stellar objects*. *Astronomical Journal*, 99, 924 – 946. <https://articles.adsabs.harvard.edu/pdf/1990AJ.....99..924B>
- Benítez-Llambay, A., & Masset, F. (2016). *The structure of protoplanetary disks*. *Astrophysical Journal*, 832(1), 105-124.
- Bernal, J., Kovetz, E. (2022). *Line-intensity mapping: theory review with a focus on star-formation lines*. *The Astronomy and Astrophysics Review*, 30(5). <https://link.springer.com/article/10.1007/s00159-022-00143-0>
- Bergardt, OI., Kusonsky, OA., Poddelsky, AI., Oinats, AV. (2022). *Self-trained artificial neural network for physical classification of ionospheric radar data*. *Advances in spaces*

- Research, 70(10), 2905-2919. <https://www-sciencedirect-com.bibliotecavirtual.unad.edu.co/science/article/pii/S0273117722006858?via%3Dihub>
- Bruderer, S., et al. (2009). *X-ray and UV radiation in protoplanetary disks*. *Astronomy & Astrophysics*, 506(2), 379-398.
- Brunt, C., et al. (2021). *Exploring machine learning for the classification of molecular clouds*. *Astronomy & Astrophysics*, 647, A28. <https://doi.org/10.1051/0004-6361/202038689>
- Buitrago, N. (2018). *Modelo estadístico para la distribución de materia en discos protoplanetarios*, Repositorio UPTC. <https://repositorio.uptc.edu.co/handle/001/2484>
- Buta, R. (2011). *The Hubble sequence of galaxies*. NASA/IPAC Extragalactic Database. <https://ned.ipac.caltech.edu/level5/Sept11/Buta/Buta10.html>
- Cánovas, H. (2018). *Discos protoplanetarios: Fábricas de planetas en acción*. *Astronomía*, (226), 22-29. https://www.sea-astronomia.es/sites/default/files/sea-astronomia_abr_2018.pdf
- Carrillo-Vargas, A. L., Pérez, L. M., & Benítez-Llambay, P. (2022). *Predicting dust gap opening in protoplanetary disks through planet-disk interactions*. *Monthly Notices of the Royal Astronomical Society*, 511(2), 2453-2465. <https://doi.org/10.1093/mnras/stab2088>
- Chaini, S., Bagul, A., Deshpande, A., Gondkar, R., Sharma, K., Vivek, M., Kembhavi, A. (2023). *Photometric identification of compact galaxies, stars, and quasars using multiple neural networks*. *Monthly Notices of the Royal Astronomical Society*, 518 (2), 3123 – 3136. <https://arxiv.org/pdf/2211.08388>
- Dieleman, S., Willett, K. W., & Dambre, J. (2015). *Rotation-invariant convolutional neural networks for galaxy morphology prediction*. *Monthly Notices of the Royal Astronomical Society*, 450(2), 1441–1459. <https://doi.org/10.1093/mnras/stv632>

- Disk Detective Collaboration. (2022). *Insights into protoplanetary disk evolution*. Astronomy & Astrophysics.
- Ethiraj, S., & Bolla, B. K. (2022). *Classification of quasars, galaxies, and stars in the mapping of the universe using multi-modal deep learning*. arXiv. <https://arxiv.org/abs/2205.10745>
- Farooq, U., et al. (2021). *Leveraging transfer learning for astronomical object classification*. TechRxiv. <https://doi.org/10.36227/techrxiv.170957049.96920652/v1>
- Fluke, C. J., & Jacobs, C. (2020). *Deep learning applications in astronomy*. Publications of the Astronomical Society of Australia, 37, e036. <https://doi.org/10.1017/pasa.2020.27>
- Forbes, J.C. Krumholz, M.R. Speagle, J.S. (2019). *Towards a radially resolved semi-analytic model for the evolution of disc galaxies tuned with machine learning*, Monthly Notices of the Royal Astronomical Society, 487(3), 3581–3606, <https://doi.org/10.1093/mnras/stz1473>
- Gao, J.L., Chen, J.Y., Wei, J.Q., Jiang, B., Luo, A.L. (2023). *Deep Multimodal Networks for M-type Star Classification with Paired Spectrum and Photometric Image*. Publications of the Astronomical Society of the Pacific, 135(1046). <https://iopscience.iop.org/article/10.1088/1538-3873/acc7ca>
- Gómez, J. (2023). *Tesis de maestría: Síntesis poblacional de Discos Protoplanetarios*. Universidad Nacional de La Plata. https://sedici.unlp.edu.ar/bitstream/handle/10915/156473/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y
- Hall, K., Yee, J., Shin, I.G., Yang, H., Zhang, J. (2024). *A new Parametrization for Finding Solutions for Microlensing Exoplanet Light Curves*. Publications of the Astronomical Society of the Pacific, 136(9).

- IEEE. (2024). *Stellar objects classification using supervised machine learning techniques*. IEEE Xplore. <https://ieeexplore.ieee.org/document/9994215>.
- IEEE. (2024). *Star Galaxy Classification Using Deep Learning*. IEEE Xplore. <https://ieeexplore.ieee.org/document/10507986>
- Jamal, S., Bloom, JS. (2020). *On Neural Architectures for Astronomical Time-series Classification with Application to Variable Stars*. *Astrophysical Journal Supplement Series*, 250(2). <https://iopscience.iop.org/article/10.3847/1538-4365/aba8ff>
- Karachentsev, I. D., et al. (2017). *Star-galaxy classification in the SDSS: Supervised learning techniques applied to photometric data*. *Monthly Notices of the Royal Astronomical Society*, 464(4), 4463–4474. <https://doi.org/10.1093/mnras/stw2657>
- Khajenabi, F., Kazrani, K., Shadmehri, M. (2017). *An Analytical Model for the Evolution of the Protoplanetary Disks*. *The Astrophysical Journal*, 841 (2), 99. <https://iopscience.iop.org/article/10.3847/1538-4357/aa6ea6>
- Kaeufer, T., Woitke, P., Min, M., Kamp, I., & Pinte, C. (2023). *Analysing the SEDs of protoplanetary disks with machine learning*. *Astronomy & Astrophysics*, 672, A30. <https://doi.org/10.1051/0004-6361/202245461>
- Kirby, A. (2023). *Exploratory Bibliometrics: using VOSViewer as a preliminary research tool*. *Publications*, 11(1), 10. <https://doi.org/10.3390/publications11010010>
- Kremer, J., Stensbo-Smidt, K., Gieseke, F., Pedersen, K. S., & Igel, C. (2017). *Big Universe, Big Data: Machine Learning and Image Analysis for Astronomy*. *IEEE Intelligent Systems, Intelligent Systems*, IEEE, *IEEE Intell. Syst*, 32(2), 16–22. <https://doi-org.bibliotecavirtual.unad.edu.co/10.1109/MIS.2017.40>

- Kumaran, S., Mandal, S., Bhattacharyya, S., Mshra, D. (2023). *Automated Classification of Chandra X-ray Point Sources using Machine Learning Methods*. Monthly Notices of the Royal Astronomical Society, 520 (4), 5065 – 5076.
<https://academic.oup.com/mnras/article/520/4/5065/7033444?login=true>
- Kvammen, A., Wickstrom, K., McKay, D., Partamies, N. (2020). *Auroral Image Classification With Deep Neural Networks*. Journal of Geophysical Research-Space Physics, 125 (10).
<https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JA027808>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. Nature, 521(7553), 436-444.
<https://doi.org/10.1038/nature14539>
- Malhotra, P., et al. (2022). *Astronomical objects classification by convolutional neural network algorithms layers*. IEEE Xplore. <https://ieeexplore.ieee.org/document/9920384>
- Mao, S., et al. (2023). *PPDONet: A fast forward problem solver for protoplanetary disks*. The Astrophysical Journal, 937(2), 165-188.
- Mandt, K.E., Mousis, O., Marty, B. et al. *Constraints from Comets on the Formation and Volatile Acquisition of the Planets and Satellites*. Space Sci Rev 197, 297–342 (2015).
<https://doi.org/10.1007/s11214-015-0161-z>
- Martins, JS., Rosenfeld, R., Sobreira, F. (2018). *Forecasts for warm dark matter from photometric galaxy surveys*. Monthly Notices of the Royal Astronomical Society, 481(1), 1290-1299. <https://academic-oup-com.bibliotecavirtual.unad.edu.co/mnras/article/481/1/1290/5078390>
- Muthukrishna, D., Parkinson, D., Tucker, B. (2019). *DASH: Deep Learning for the Automated Spectral Classification of Supernovae and Their Hosts*. The Astrophysical Journal, 885(85). <https://iopscience.iop.org/article/10.3847/1538-4357/ab48f4>

- O'Dell, C., Wen, Z. (1994). *Postrefurbishment Mission Hubble Space Telescope Images of the Core of the Orion Nebula: Proplyds, Herbig-Haro Objects, and Measurements of a Circumstellar Disk*. *Astrophysical Journal*, 436, 194 – 202.
<https://articles.adsabs.harvard.edu/pdf/1994ApJ...436..194O>
- Ohta, Y., et al. (2018). *Supervised machine learning techniques for the detection of exoplanets from light curves*. *Astrophysical Journal*, 852(2), 123. <https://doi.org/10.3847/1538-4357/aaa8f9>
- Orlando, E., Anugu, N., Hsieh, C., Karnath, N. (2024, enero 8). *AAS 243 Press Conference: stars, protostars & more clouds [Conferencia de prensa]*. American Astronomical Society's Press Office. <https://www.youtube.com/watch?v=9zUvGTXmN24>
- Padovani, M., et al. (2018). *Cosmic ray ionization in protoplanetary disks*. *Astronomy & Astrophysics*, 615, A61.
- Pawlik, M. (2014). *Quantifying galaxy morphology*. *Astronomy & Geophysics*, 55(6), 6.13-6.16.
<https://doi.org/10.1093/astrogeo/atu246>
- Pontificia Universidad Católica de Chile. (n.d.). *Ciencia astronómica se centra en inteligencia artificial y nuevas tecnologías*. <https://www.uc.cl/noticias/ciencia-astronomica-se-centra-en-inteligencia-artificial-y-nuevas-tecnologias/>
- Qu, H., Sako, M., Möller, A., Doux, C. (2021). *SCONE: Supernova Classification with a Convolutional Neural Network*. Arxiv. <https://arxiv.org/abs/2106.04370>
- Ravanbakhsh, S., et al. (2016). *Galaxy morphology and cosmological model discrimination using machine learning*. *Monthly Notices of the Royal Astronomical Society*, 462(3), 3464-3481. <https://doi.org/10.1093/mnras/stw1918>

ResearchGate. (2023). *Study on protoplanetary disk evolution and chemistry*.

<https://www.researchgate.net/publication/333641812>.

Reza, M. (2021). *Galaxy morphology classification using automated machine learning*.

Astronomy and Computing, 37, 1 – 11. <https://doi.org/10.1016/j.ascom.2021.100492>

Rodríguez-Feliciano, B., Serna, J., García-Varela, A., Hernández, J., Cabrera-García, JF. (2023).

Machine-learning Morphological Classification of TESS Light Curves of T Tauri Stars.

Astronomical Journal, 166 (5). [https://iopscience.iop.org/article/10.3847/1538-](https://iopscience.iop.org/article/10.3847/1538-3881/acf865)

[3881/acf865](https://iopscience.iop.org/article/10.3847/1538-3881/acf865)

Rodríguez, J. V., Rodríguez-Rodríguez, I., & Woo, W. L. (2022). *On the application of machine learning in astronomy and astrophysics: A text-mining-based scientometric analysis*.

Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(5), e1476.

Roy, N., Napolitano, N., La Barbera, F., Tortora, C., Getman, F., Radovich, M., Capaccioli, M.,

Brescia, M., Cavuoti, S., Longo, G., Raj, M., Puddu, E., Covone, G., Amaro, V., Velluci,

C., Grado, A., Kuijken, K., Versoes Kleijn, G., Valentijn, E. (2018). *Evolution of Galaxy*

size-stellar mass relation from the Kilo-degree Survey. *Monthly Notices of the Royal*

Astronomical Society, 480(1). <https://doi.org/10.1093/mnras/sty1917>

Sabeesh, E., & Bolla, B. K. (2020). *Leveraging deep learning for the classification of*

astronomical objects. *Astronomy & Astrophysics*, 642, A50.

<https://doi.org/10.1051/0004-6361/202038680>

Schmelzle, J., et al. (2017). *Cosmological model discrimination with Deep Learning*.

arXiv:1707.05327. <https://arxiv.org/abs/1707.05327>

- Sen, S., Agarwal, S., Chakraborty, P., & Singh, K. P. (2022). *Astronomical big data processing using machine learning: A comprehensive review*. *Experimental Astronomy*, 53(1), 1-43.
<https://doi.org/10.1007/s10686-021-09827-4>
- Shang, ZY., Yao, ZH., Liu, J., Xu, LL., Xu, Y., Zhang, BZ., Guo, RL., Wei, Y. (2023). *Automated Classification of Auroral Images with Deep Neural Networks*. *Universe*, 9(2).
<https://www.mdpi.com/2218-1997/9/2/96>
- Smith, B., Terrile, R. (1994). *A circumstellar disk around β Pictoris*. *Science*, 4681(226), 1421 – 1424. <https://www.science.org/doi/10.1126/science.226.4681.1421>
- Strom, K., Strom, E., Edwards, S., Cabrit, S., Skrutskie, M. (1989). *Circumstellar material associated with solar-type pre-main-sequence stars: a possible constraint on the timescale for planet building*. *Astronomical Journal*, 97, 1454 – 1470.
<https://adsabs.harvard.edu/full/1989AJ.....97.1451S>
- Takahashi, S., Muto, T. (2018). *Structure Formation in a Young Protoplanetary Disk by a Magnetic Disk Wind*. *Astrophysical Journal*, 865(2), 102.
<https://arxiv.org/pdf/1808.10220>
- Terry, J., Hall, C., Abreau, S., & Gleyzer, S. V. (2022). *Locating Hidden Exoplanets in ALMA Data Using Machine Learning*. *Astrophysical Journal/The Astrophysical Journal*, 941(2), 192. <https://doi.org/10.3847/1538-4357/aca477>
- Toba, Y., et al. (2019). *Quasar classification using supervised learning techniques on SDSS data*. *Publications of the Astronomical Society of Japan*, 71(4), 80.
<https://doi.org/10.1093/pasj/psz101>

- Universidad de Hertfordshire. (s.f.). *The formation and evolution of galaxies*. University of Hertfordshire. <https://www.herts.ac.uk/research/centres/car/extragalactic/the-formation-and-evolution-of-galaxies>
- Trierweiler, I. (2024). *Machine learning in protoplanetary disk chemistry*. Astrobites.
- Vaia (2023). *Light curve analysis: Definition & techniques*. Vaia. www.vaia.com
- Vanderplas, J.T., Connolly, A.J., Ivezić, Ž., & Gray, A. (2012). *Introduction to astroML: Machine learning for astrophysics*. In Conference on Intelligent Data Understanding (CIDU) (pp. 47-54). IEEE. <https://doi.org/10.1109/CIDU.2012.6382200>
- Vujcic, D., Pavlovic, R., Milosevic, D., Djordevic, B., Randjic, S., Stojic, D. (2020). *Classification of asteroid families with artificial neural networks*. Serbian Astronomical Journal, 201, 39-48. <https://doiserbia.nb.rs/Article.aspx?ID=1450-698X2001039V>
- Wang, K., Guo, P., Yu, FS., Duan, LZ., Wang, YP., Du, H. (2018). *Computational Intelligence in Astronomy: A survey*. International Journal of Computational Intelligence Systems, 11(1), 575-590. <https://doi.org/10.2991/ijcis.11.1.43>
- Wang, YT., Zhao, GB. (2021). *Cosmology with galaxy redshift surveys: Probing dark energy*. Chinese Science Bulletin-Chinese, 66(11), 1336-1345. <https://dx-doi-org.bibliotecavirtual.unad.edu.co/10.1360%2FTB-2021-0071>
- Weintraub, D., Sandell, G., Duncan, W. (1989). *Submillimeter measurements of T Tauri and FU Orionis Stars*. Astrophysical Journal Letters, 340, L69 – L72. <https://adsabs.harvard.edu/full/1989ApJ...340L..69W>
- Williams, J., Cieza, L. (2010). *Protoplanetary disks and their evolution*. Institute for Astronomy, University of Hawaii, 49, 1056. <https://ar5iv.labs.arxiv.org/html/1103.0556>

- Xu, D., Zhu, Y. (2024). *Surveying image segmentation approaches in astronomy*. *Astronomy and Computing*, 48. <https://doi.org/10.1016/j.ascom.2024.100838>
- Xu, YJ., Huang, WR., Deng, H., Mei, Y., Wang, F. (2020). *A Hybrid Method of Accurate Classification for Blazars of Uncertain Type in Fermi-LAT Catalogs*. *Astrophysical Journal*, 895(2). <https://iopscience.iop.org/article/10.3847/1538-4357/ab8ae3>
- Yuan, H., Liu, M., Yang, ZC., Bu, YD., Yi, ZP., Kong, XM., Wu, TM., Zhang, R. (2023). *Red Giants Search Method Based on Convolutional Neural Networks*. *Astronomical Journal*, 166(6). <https://iopscience.iop.org/article/10.3847/1538-3881/ad0297>
- Zhang, C., & Zhao, Y. (2015). *Automatic classification of astronomical objects using deep learning*. arXiv. <https://arxiv.org/abs/1511.06389>
- Zhang, JY., Zhang, YX., Kang, ZH., Li, CH., Tao, YH., Zhao, YH., Wu, XB. (2023). *A novel approach for variable star classification based on imbalanced learning*. *Publications of the Astronomical Society of Australia*, 40. <https://www-cambridge-org.bibliotecavirtual.unad.edu.co/core/services/aop-cambridge-core/content/view/723B802AEA0E9BD24D747BDDA6B2D525/S1323358023000358a.pdf/a-novel-approach-for-variable-star-classification-based-on-imbalanced-learning.pdf>
- Zhao, H., Sun, TY., Yu, SX. (2024). *Deep Neural Network Closed-loop with Raw Data for Optical Resident Space Object Detection*. *Research in Astronomy and Astrophysics*, 24(11). <https://iopscience.iop.org/article/10.1088/1674-4527/ad849a>