



Universidad Cenfotec

Maestría en tecnología de bases de datos

Documento final de Proyecto de Investigación Aplicada 2

Estimación del impacto del tránsito de los satélites en la observación astronómica en el territorio centroamericano utilizando una arquitectura basada en la nube.

Campos Miranda Oscar Mario

Chinchilla Elizondo Guillermo E.

Marzo, 2021

Declaratoria de derecho de autor

Se declara que el presente proyecto de investigación fue realizado por los autores Oscar Mario Campos Miranda y Guillermo E. Chichilla Elizondo, utilizando las fuentes bibliográficas y trabajos citados, cuales tienen su respectiva referencia, respetando los derechos de autor de dichos trabajos.

Se autoriza la reproducción total o parcial de este trabajo, para ser usados como referencia de trabajos futuros de tipo académico y científico, en este caso, se solicita incorporar la referencia de este trabajo respetando los derechos de los autores.

Agradecimientos

Queremos agradecer a nuestras familias quienes brindaron apoyo y acompañamiento tanto en el desarrollo del presente trabajo como en el programa de la maestría, brindándonos la motivación para finalizarla de manera exitosa.

También queremos agradecer a M. Sc. Brisa Terezón como directora del observatorio Micro Macro de la Universidad Don Bosco quien nos brindó el espacio para presentar los resultados de la investigación, así como excelente retroalimentación durante el proceso.

Tabla de Contenido

Abstract	1
Capítulo 1. Introducción	2
1.1 Generalidades	2
1.2 Antecedentes del Problema	2
1.3 Definición y Descripción del Problema	3
1.4 Justificación	5
1.5 Viabilidad	6
1.5.1 Punto de Vista Técnico.	6
1.5.2 Punto de Vista Operativo.	6
1.5.3 Punto de Vista Económico.	6
1.6 Objetivos	7
1.6.1 Objetivo General.	7
1.6.2 Objetivos Específicos.	7
1.7 Alcances y Limitaciones	8
1.7.1 Alcances.	8
1.7.2 Limitaciones.	8
1.8 Marco de Referencia Organizacional y Socioeconómico	8
1.9 Estado de la Cuestión	10
Capítulo 2. Marco Teórico o Conceptual	27
Capítulo 3. Marco Metodológico	46
3.1 Tipo de Investigación	46
3.2 Alcance Investigativo	46
3.3 Enfoque	47
3.4 Diseño	49
3.5 Población y Muestreo	50
3.6 Instrumentos de Recolección de Datos	50

3.7 Técnicas de Análisis de Información	51
Capítulo 4. Análisis del Diagnóstico.....	53
4.1 Catálogo de satélites	53
4.2 Elementos de dos líneas (TLE)	55
4.3 Magnitudes estándar	56
Capítulo 5. Propuesta de Solución	58
5.1 Lenguaje de programación y bibliotecas a utilizar	58
5.2 Infraestructura de hardware y tecnologías de almacenamiento	58
5.3 Extracción, transformación y carga de los datos.....	59
5.4 Presentación y análisis de resultados	83
5.4.1 Detalle de la interfaz	85
5.4.2 Modelo futuro.....	87
5.4.3 Predicción de tránsitos	88
5.4.4 Análisis de Resultados.....	89
5.5 Instrucciones de Implementación del sistema	93
Capítulo 6. Conclusiones y Recomendaciones	104
6.1 Conclusiones	104
6.2 Recomendaciones	106
Capítulo 7. Reflexiones Finales	108
Capítulo 8. Trabajos a Futuro {Líneas Futuras de Investigación}.....	110
Referencias	111
Apéndices.....	117

Lista de figuras

Figura 1: Primer ejemplo de interferencia de satélite en imágenes de alta exposición. Fuente: https://www.space.com/astronomy-group-worries-about-starlink-science-interference.html	3
Figura 2: Segundo ejemplo de interferencia de satélite en imágenes de alta exposición. Fuente: https://www.vox.com/science-and-health/2020/1/7/21003272/space-x-starlink-astronomy-light-pollution	4
Figura 3: Diagrama del proceso de revisión de literatura. Fuente: Elaboración propia	14
Figura 4: Nube de conceptos. Fuente: Elaboración propia.....	27
Figura 5: Mapa conceptual. Fuente: Elaboración propia	27
Figura 6: Ubicación de un objeto utilizando el azimut. Adaptada al español del sitio web heavens-above.com	29
Figura 7: Ilustración del cénit. Adaptada al español de https://www.pveducation.org/pvcdrom/properties-of-sunlight/elevation-angle ..	30
Figura 8: Ejemplos de magnitud aparente. Tomada de https://www.astronomia-iniciacion.com/magnitudes.html	31
Figura 9: Visualización del reflejo satelital Tomada de https://www.high-iso.de/blog/tutorials/iridium-flares-fotografieren/ . Video convertido a imágenes para una mejor interpretación.	34
Figura 10: Ejemplo de un elemento de dos líneas. Adaptada al español de https://spaceflight.nasa.gov/realdata/sightings/SSapplications/Post/JavaSSOP/SOP_Help/tle_def.html	35
Figura 11: Representación ontológica de la investigación. Fuede: elaboración propia	48

Figura 12: Ejemplo de magnitud estándar tomada de: http://www.satflare.com/search.asp?sat=STARLINK-41	51
Figura 13: Diagrama de flujo del proceso de obtención y análisis de la información. Fuente: Elaboración propia	52
Figura 14: Desglose de costos mensuales de los recursos de infraestructura. Fuente: Elaboración propia	59
Figura 15: Encabezado de la solicitud HTTP para la extracción del catálogo de satélites. Fuente: Elaboración propia	62
Figura 16: Código para extracción del catálogo de satélites. Fuente: Elaboración propia	62
Figura 17: Código para extracción de TLEs	63
Figura 18: Estructura del proceso de carga del catálogo de satélites. Fuente: Elaboración propia	66
Figura 19: Vista general del proceso de carga de los elementos de dos líneas. Fuente: Elaboración propia	67
Figura 20: Vista del proceso de ejecución de cada lote de los elementos de dos líneas. Fuente: Elaboración propia	68
Figura 21: Estructura del proceso de carga de la magnitud estándar. Fuente: Elaboración propia	68
Figura 22: Proceso del cálculo de los tránsitos satelitales. Fuente: Elaboración propia	70
Figura 23: Vista general del proceso de cálculo de los tránsitos. Fuente: Elaboración propia	71
Figura 24: Procesamiento detallado del cálculo de los tránsitos por país. Fuente: Elaboración propia	71

Figura 25: Flujo de procesos modificado. Fuente: Elaboración propia	74
Figura 26: Escala para la clasificación del impacto en la observación astronómica. Fuente: Elaboración propia	75
Figura 27: Ejemplo de los datos previo a la selección del tránsito de mayor impacto. Fuente: Elaboración propia	76
Figura 28: Cantidad de satélites por año. Fuente: Elaboración propia	77
Figura 29: Resultados del modelo de predicción utilizando toda la información disponible. Fuente: Elaboración propia	79
Figura 30: Resultados del modelo de predicción utilizando información de los últimos 5 años. Fuente: Elaboración propia.	80
Figura 31: Código utilizado para creación del modelo de predicción. Fuente: Elaboración propia.....	81
Figura 32: Código de evaluación del modelo de predicción. Fuente: Elaboración propia.....	82
Figura 33: Visualización de clasificación de impacto visual. Fuente: Elaboración propia	85
Figura 34: Filtro de fecha. Fuente: Elaboración propia.....	85
Figura 35: Filtro de tiempo de exposición. Fuente: Elaboración propia	86
Figura 36: Filtro de ciudad. Fuente: Elaboración propia	86
Figura 37: Filtro de impacto. Fuente: Elaboración propia	86
Figura 38: Filtro de hora Fuente: Elaboración propia	86
Figura 39: Tarjeta de cantidad de satélites activos Fuente: Elaboración propia	87
Figura 40: Visualización de crecimiento anual de satélites activos Fuente: Elaboración propia.....	87

Figura 41: Predicción de cantidad de satélites para el próximo año. Fuente: Elaboración propia.....	88
Figura 42: Predicción de cantidad de satélites para el próximo año. Fuente: Elaboración propia.....	89
Figura 43: Grafico de clasificación de tránsitos por hora utilizando la configuración de 1000 segundos de exposición. Fuente: Elaboración propia	90
Figura 44: Grafico de clasificación de tránsitos por hora utilizando la configuración de 1000 segundos de exposición para El Salvador. Fuente: Elaboración propia	90
Figura 45: Visualización de predicción del crecimiento en la cantidad de satélites para el año 2021. Fuente: Elaboración propia.	92
Figura 46: Cálculo de aumento de tránsitos basado en porcentaje de aumento pronosticado. Fuente: Elaboración propia.....	92
Figura 47: Ejemplo de comando de instalación de dependencias de azure en PowerShell. Fuente: Elaboración propia	94
Figura 48: Menú contextual de carpeta en Windows. Fuente: Elaboración propia ..	94
Figura 49: Llamada al script de despliegue inicial. Fuente: Elaboración propia.....	95
Figura 50: Solicitud de credenciales por parte del script de despliegue. Fuente: Elaboración propia.....	95
Figura 51: Búsqueda del grupo de recursos implementado. Fuente: Elaboración propia.....	96
Figura 52: Identificación de la cuenta Batch en el grupo de recursos. Fuente: Elaboración propia.....	97
Figura 53: Ejemplo de cómo agregar una aplicación en la cuenta Batch parte1. Fuente: Elaboración propia	98

Figura 54: Ejemplo de cómo agregar una aplicación en la cuenta Batch parte2. Fuente: Elaboración propia	98
Figura 55: Vista final de las cuatro aplicaciones creadas. Fuente: Elaboración propia	99
Figura 56: Guía para enlazar las aplicaciones al pool parte 1. Fuente: Elaboración propia.....	100
Figura 57: Guía para enlazar las aplicaciones al pool parte 2. Fuente: Elaboración propia.....	100
Figura 58: Ejemplo de conexión a base de datos utilizando SSMS. Fuente: Elaboración propia.....	101
Figura 59: Ejemplo de ejecución de script utilizando SSMS. Fuente: Elaboración propia	101

Lista de tablas

Tabla 1: Palabras a incluir en la revisión sistemática	11
Tabla 2: Formulario de extracción de información. Fuente: Elaboración propia.	16
Tabla 3: Primera fuente literaria investigativa. Información obtenida de McDowell (2020)	17
Tabla 4: Segunda fuente literaria investigativa. Información obtenida de Seitzer (2020)	19
Tabla 5: Tercera fuente literaria investigativa. Información obtenida de Hainaut (2020)	20
Tabla 6: Cuarta fuente literaria investigativa. Información obtenida de Gallozi (2020)	22
Tabla 7: Quinta fuente literaria investigativa. Información obtenida de Ren (2019) .	23

Tabla 8: Sexta fuente literaria investigativa. Información obtenida de Bai (2019)....	25
Tabla 9: Presentación de resultados de revisión sistemática	25
Tabla 10: Escala para la clasificación del impacto en la observación astronómica. Fuente: Elaboración propia	49
Tabla 11: Determinación de clasificación por nota ponderada. Fuente: Elaboración propia	49
Tabla 12: Diccionario de datos del catálogo de satélites. Fuente: Elaboración propia	54
Tabla 13: Detalle de objetos rastreados por SpaceTrack.org. Fuente: Elaboración propia	54
Tabla 14: Atributos en conjunto de datos de TLE. Fuente: Elaboración propia	56
Tabla 15: Estructura de tabla para almacenar los tránsitos. Fuente: Elaboración propia	61
Tabla 16: Desglose del URL para retornar el catálogo de satélites. Fuente: Elaboración propia.....	62
Tabla 17: Desglose del URL para retornar los elementos de dos líneas. Fuente: Elaboración propia.....	63
Tabla 18: Distribución de satélites de acuerdo con la antigüedad de TLE.....	64
Tabla 19: Descripción del conjunto de datos de Satflare.com. Fuente: Elaboración propia.....	65
Tabla 20: Descripción de la forma final de los datos de magnitud satelital. Fuente: Elaboración propia.....	66
Tabla 21: Conteo de satélites con magnitud intrínseca. Fuente elaboración propia	73
Tabla 22: Representación de la preparación del conjunto de datos. Fuente: Elaboración propia.....	78

Tabla 23: Representación de la predicción de datos. Fuente: Elaboración propia. .	79
Tabla 24: Resultados del modelo de diagnóstico. Fuente: Elaboración propia.....	83
Tabla 25: Objetos de base de datos necesarios almacenar y calcular la información. Fuente: Elaboración propia	103

Abstract

El proyecto Starlink de la empresa SpaceX, pretende ofrecer internet satelital, para lo cual requiere establecer una mega constelación de satélites artificiales.

Este hecho ha levantado alertas en la comunidad científica, quienes consideran que este incremento en la cantidad de satélites generará que las observaciones sean afectadas por la suma de tránsitos satelitales que obstruirían la visibilidad.

El análisis realizado a los datos demuestra que actualmente, el mayor impacto de los tránsitos satelitales se da en dos momentos del día, cerca del amanecer y poco después del atardecer. Además, se obtiene pronóstico de crecimiento en la cantidad de satélites de entre 12% y 13% para finales del 2021 con respecto al mismo periodo de 2020.

Este estudio también muestra de manera detallada el diseño e implementación del sistema de inteligencia de negocios desarrollado para realizar los cálculos y monitorear la actividad en la región, con el fin de mantener las estadísticas actualizadas.

Capítulo 1. Introducción

1.1 Generalidades

Para efectos del presente trabajo la definición de “impacto en las observaciones” será dada por la cantidad y clasificación de satélites que puedan presentar algún nivel de interferencia en la toma de imágenes por parte de los observatorios en un momento dado.

Por lo tanto, la creación de los modelos es de aplicación teórica, y todas las estimaciones realizadas en este trabajo no han sido emparejadas con observaciones de campo. Así pues, la cuantificación del impacto se realiza asumiendo varios parámetros que en la práctica podrían mostrar variabilidad por diferentes factores, como la precisión del cálculo del reflejo, metodología de observación, entre otras.

La información utilizada para el cálculo de las órbitas de los satélites es de acceso público, pero a pesar de ello, el autor no está en la capacidad de compartir los datos obtenidos para la realización de este trabajo, ya que existe una restricción en las condiciones de uso en cuanto al reenvío a terceros. Por lo que en caso que el lector necesitare de estos, debe obtenerla directamente del sitio web Space-Track.org.

1.2 Antecedentes del Problema

Existen análisis preliminares, del impacto que el despliegue de las constelaciones satelitales, causa en diferentes observatorios, ubicados en distintas latitudes terrestres. Estos estudios, han mostrado la situación en un momento determinado, enfocado a observatorios o situaciones particulares.

Ninguno de estos análisis, muestra el impacto directo sobre territorio centroamericano, utilizando técnicas de minería de datos y poniendo a disposición un reporte interactivo, que se mantenga actualizado y que pueda ser utilizado por observatorios profesionales o universidades, para el estudio de la situación actual y futura.

Adicionalmente se cuenta con varias aplicaciones que proveen la ubicación y pronóstico del paso de satélites en tiempo real, no obstante, estas aplicaciones no cuentan con un análisis del impacto visual acumulado o totalizado (actual e histórico) en un momento y ubicación determinada.

Dadas las condiciones actuales los investigadores podrían ver afectados sus estudios, debido a que, se proyecta que la interferencia de los satélites sea mayor conforme pase el tiempo. Esta situación se traduce en, pérdidas de horas e información relevante, que imposibilita la obtención de resultados para el estudio, desgasta los recursos de este, e implica destinar más tiempo para el análisis.

Es, precisamente, la intención de la presente investigación, solventar la necesidad de brindar un reporte que calcule el impacto visual de las constelaciones satelitales, en la observación astronómica centroamericana en un momento determinado y así ayude a los usuarios de este reporte a hacer mejores inversiones y estar atentos a los cambios en el panorama regional.

1.3 Definición y Descripción del Problema

El primer satélite artificial puesto en órbita fue el Sputnik, en 1957, construido por la Unión Soviética. Desde entonces, la cantidad de satélites, desechos espaciales producidos por colisiones y restos de otros lanzamientos ha ido en aumento, causando que cada vez haya más cuerpos contaminantes en órbita.

Estos objetos que orbitan la tierra están hechos de diferentes materiales que reflejan la luz del sol hacia la tierra con diferente nivel de intensidad, lo que causa que los instrumentos más sensibles a la luz los detecten, afectando así, las observaciones como se puede ejemplificar en la Figura 1 y Figura 2.

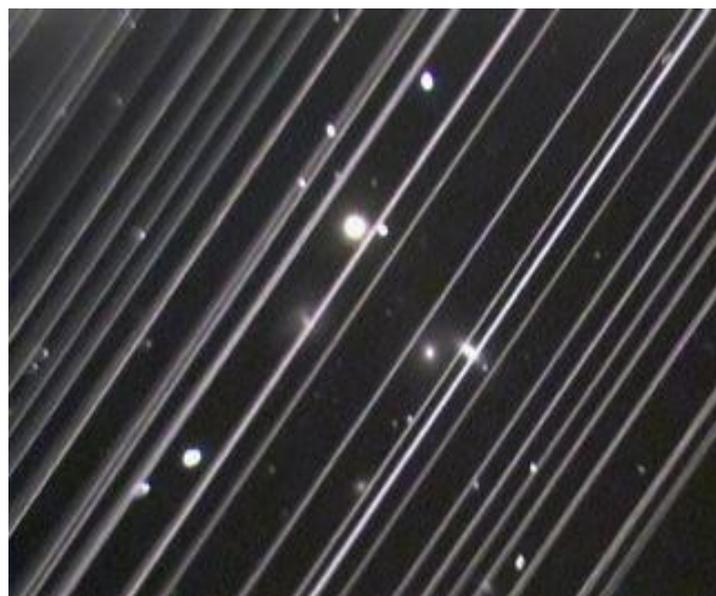


Figura 1: Primer ejemplo de interferencia de satélite en imágenes de alta exposición. Fuente: <https://www.space.com/astronomy-group-worries-about-starlink-science-interference.html>

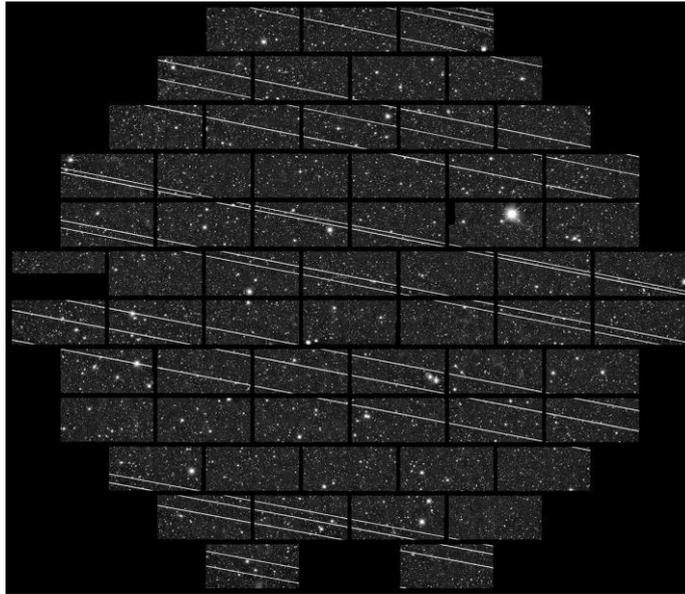


Figura 2: Segundo ejemplo de interferencia de satélite en imágenes de alta exposición. Fuente: <https://www.vox.com/science-and-health/2020/1/7/21003272/space-x-starlink-astronomy-light-pollution>

En años anteriores, estos efectos se han podido contrarrestar mediante la planificación de las observaciones, utilizando herramientas de rastreo satelital. Programando la captura de imágenes en momentos donde no se esperen satélites cruzando el cielo del observador en el área por estudiar.

Sin embargo, con el reciente despliegue de la constelación de satélites de Starlink de SpaceX, la comunidad científica alzó la voz indicando que existe un grave riesgo, en cuanto a contaminación visual, que podría aumentar el problema en varias órdenes de magnitud. Según Gallozi (2020) podría ser de un 300% de incremento tan solo tomando en cuenta la flota Starlink los cuales ya tienen lanzamientos programados para los próximos años, sin tomar en cuenta futuras misiones de otras empresas como Amazon.

Estas constelaciones de satélites, más los objetos ya existentes orbitando la atmósfera terrestre, ponen en riesgo las operaciones de los observatorios, los cuales necesitan condiciones óptimas de oscuridad, sin interferencia para propiciar que las imágenes obtenidas sean utilizables.

Es por esta razón, que se propone en esta investigación, proveer un sistema que permita calcular el impacto visual, de acuerdo al país, tiempo de observación y pronóstico; que facilite el aprovechamiento adecuado de los recursos de los observatorios locales, además de la concientización sobre la situación actual.

1.4 Justificación

En la observación astronómica profesional por medios ópticos, la calidad de las imágenes es de suma importancia para el análisis de los objetos del espacio profundo, en la medida de lo posible, se requiere de condiciones de oscuridad total sin ninguna obstrucción. Cualquier interferencia o problema técnico que suceda durante la captura de dichas imágenes supone la pérdida de horas de trabajo y utilización de recursos. Estudios preliminares para observatorios europeos (Hainaut & Williams 2020) indican que la afectación en las observaciones podría ser tan grande como el 30% de las imágenes de alta exposición durante el amanecer y hasta el 50% durante el anochecer.

Es por esto que, esta propuesta está destinada a ayudar a que se puedan programar con mayor eficacia las observaciones, de modo que se aprovechen de mejor forma los recursos utilizados en las observaciones. Así también evitar inversiones en equipos que no puedan ser aprovechados de una manera efectiva mediante el análisis de la situación en la región.

Según la comunidad científica, con el paso de los años, tendremos un aumento exponencial de satélites en órbita, que agrava el impacto de los mismos en la observación.

Dado lo anterior, la presente investigación tiene como fin, la realización de un reporte interactivo, que pueda mantener la información más reciente, sobre la ubicación de los satélites. El cual facilite, a los observatorios la identificación, de manera oportuna, ágil y rápida; la cantidad de satélites en órbita, que puedan presentar un nivel de afectación en las observaciones. Así como estadísticas generales históricas sobre la hora y época del año con mayor afectación. Esta herramienta, les ayudará a identificar, las horas adecuadas para la observación y los mantendrá informados de los cambios en el panorama actual.

De la misma forma, está orientado a optimizar el uso de los recursos del observatorio, así como las inversiones futuras, realizadas en capital humano e infraestructura tecnológica, la cual asciende a cientos de miles de dólares.

1.5 Viabilidad

1.5.1 Punto de Vista Técnico.

Para el presente trabajo, se requiere conocimiento técnico en tecnologías de minería de datos como lo son manejo de tecnologías de bases de datos SQL y NoSQL, análisis de series de tiempo, predicción utilizando técnicas de regresión lineal, agrupación, plataformas de procesamiento en la nube, por ejemplo, Azure y herramientas de reportería como Power BI. Dado que los autores poseen el conocimiento en dichas áreas, y han trabajado con la aplicación de estas tecnologías, se cuenta con la experiencia técnica para desarrollar el sistema adecuado a las necesidades de la institución.

1.5.2 Punto de Vista Operativo.

La investigación y desarrollo del reporte para la institución, requiere la consulta y extracción de datos que brinda la base de datos pública del sitio <https://www.space-track.org>, por lo que, no genera una interrupción en el desempeño general del observatorio. Sin embargo, se coordinará con el personal, mediante sesiones virtuales, con el fin de ampliar sobre las necesidades que requieren solventar. Dado que la investigación está dirigida a brindar una herramienta que beneficia la observación astronómica, se cuenta con la anuencia e interés del observatorio a participar de dichas sesiones.

1.5.3 Punto de Vista Económico.

La investigación utiliza tecnologías en la nube. Para realizar la misma, los autores, cuentan con acceso a dicha tecnología, con licencia personal. Por lo anterior, se cuenta con la viabilidad económica para llevar a cabo este estudio; los investigadores asumirán el costo del mismo. El observatorio no deberá cubrir ninguno de estos costos durante la investigación.

En cuanto a recurso humano, el trabajo será realizado por dos ingenieros en sistemas informáticos. Según el colegio de ciencias económicas costarricense el costo de la hora por servicios profesionales se definió en ₡28.293,69. Dicho lo anterior, tomando en cuenta que se estima que el presente trabajo se extienda por alrededor de 30 semanas y que por cada semana se estima que se dediquen 30 horas de trabajo, esto llevaría a un total de 900 horas. Por cuanto el costo teórico equivale a ₡25.464.321,00, el cual será asumido por los investigadores.

1.6 Objetivos

Para el presente trabajo se utiliza la taxonomía de Bloom de 1956 ya que su estructura jerárquica facilita la creación de conocimiento de manera ordenada y comprensible para el lector.

1.6.1 Objetivo General.

Estimar el impacto del tránsito de los satélites en la observación astronómica en el territorio centroamericano, utilizando una arquitectura basada en la nube para el mejor aprovechamiento de los recursos.

1.6.2 Objetivos Específicos.

- Adquirir los datos necesarios para el cálculo del tránsito satelital utilizando la base de datos Space-Track de la fuerza aérea de los Estados Unidos.
- Transformar los datos para su correcto almacenamiento en una estructura de base de datos mediante un proceso automatizado de extracción, transformación y carga.
- Emplear algoritmos de ubicación satelital con el fin de predecir la posición en el cielo del observador.
- Clasificar los satélites de acuerdo a su distancia y magnitud aparente (brillo) promedio, para asignar el peso de su posible impacto en las observaciones astronómicas.
- Confeccionar un reporte que muestre el recuento de los satélites que presentan algún grado de interferencia en la observación y su clasificación en un momento determinado.
- Diseñar un reporte que muestre la estimación de la cantidad de satélites proyectado a un año.

1.7 Alcances y Limitaciones

1.7.1 Alcances.

El alcance de este proyecto es la elaboración de una base de datos que almacene la información correspondiente a los datos públicos de los satélites que se encuentren en órbita. Incluyendo el catálogo de satélites, la información TLE (Two line element), la cual es la información utilizada para el rastreo y posicionamiento de los mismos.

Se entrega también un sistema de extracción, transformación y carga de datos, junto con su manual de implementación.

Además, se realiza un reporte interactivo, el cual se conecta a la base de datos para mostrar los resultados del procedimiento de predicción y cálculo de la órbita, basado en la información más reciente existente y los parámetros incluidos por el usuario.

1.7.2 Limitaciones.

Este trabajo no incluye un manual de mantenimiento, monitoreo, ni optimización de la base de datos, procesos de carga ni lectura de información.

Esta investigación, tampoco contiene documentación sobre el funcionamiento o arquitectura de los lenguajes de programación, motores de bases de datos, herramientas ni licencias de software utilizadas para la creación del sistema indicado.

La investigación tampoco presenta un modelo de inteligencia artificial para completar o inferir datos acerca de la magnitud intrínseca, ya que no se tuvo acceso a una base de datos que presentara características que se pudieran utilizar para tal fin.

1.8 Marco de Referencia Organizacional y Socioeconómico

1.8.1 Historia.

El Observatorio Micro Macro (OMM), está ubicado en El Salvador y pertenece a la Universidad Don Bosco. Se considera, el mayor observatorio de Centroamérica y alberga al telescopio Ritchey-Chrétien, que fue creado por la empresa alemana Alluna Optics. Este telescopio mide 50 cm de diámetro, posee una montura ecuatorial, y su elevación es de 13 grados. Su control y el del domo, se da de

manera remota por un software. Por lo que, es utilizado para la astrografía y el estudio astrológico.

Está equipado, además, con cámaras CCD utilizadas en astronomía para transmitir imágenes del telescopio en tiempo real, tales como la luna, planetas, galaxias, cúmulos globulares y abiertos, con el objetivo de realizar investigaciones en este campo.

Fue inaugurado en 2019 y nace del convenio con la Fundación Karlheinz Wolfgang de Alemania y la Universidad Don Bosco de El Salvador. Este proyecto consta de tres etapas: la primera, se trata de la infraestructura física del observatorio, la segunda es el equipamiento tecnológico necesario para su funcionamiento y la tercera, se refiere al aspecto académico-científico propio de un proyecto universitario que dará un impulso a nuevas carreras y profesiones.

1.8.2 Tipo de Negocio y Mercado Meta.

La institución nace en el seno universitario Don Bosco, por lo que está dirigida a la comunidad profesional e investigativa. Sin embargo, dicha universidad posee un enfoque social, por lo que busca acercar a toda persona interesada en conocer la bastedad del universo, tal como lo citan en su página oficial:

“La visita al OMM irá más allá, ya que, a través de este escenario práctico, cada persona experimentará de una manera vivencial, la relación entre las acciones de un ser humano (a nivel micro) y el impacto de estas en la conformación de la familia, la sociedad y el país (a nivel macro) como resultado de la suma de varios aportes individuales.”

(OMM, 2019)

1.8.3 Misión, Visión y Valores.

Tomando en cuenta lo publicado por parte del observatorio en su página oficial, en el 2019, la visión y la misión de esta institución son:

Misión del OMM

“Promover experiencias científicas para potenciar las competencias humanas, científicas y profesionales de los miembros de la comunidad universitaria y ponerlas a disposición de las personas de todas las edades, dentro y fuera

del país, usando la tecnología como herramienta para conectar nuestras acciones con el universo.”

(OMM, 2019)

Visión del OMM

“Ser un referente de divulgación de las ciencias y la tecnología, especialmente de la astronomía y sus áreas afines, siendo el único observatorio astronómico instalado en una universidad para ser usado como herramienta didáctica para potenciar el desarrollo de la persona.”

(OMM, 2019)

1.8.4 Políticas Institucionales.

Es necesario apegarse al protocolo de investigaciones con el que cuenta la institución.

1.9 Estado de la Cuestión

1.9.1 Planificación de la revisión

1.9.1.1 Formulación de la pregunta

Identificar estudios que provean información acerca de la afectación provocada por los satélites en la observación astronómica.

1.9.1.2 Amplitud y calidad de la pregunta

- Problema: El despliegue de redes de satélites produce una contaminación lumínica en los cielos nocturnos cada vez mayor. Esta tendencia va a seguir al alza debido a que el reflejo de luz de los objetos que orbitan la tierra causa que las fotografías y mediciones astronómicas se vean afectadas, teniendo así resultados no deseados en la captura de fotografías y ondas de radio. De esto se desprende el problema a investigar, que es estimar el impacto teórico que el reflejo de los satélites tendría sobre el OMM.
- Pregunta de investigación: ¿Cuáles soluciones, de las diversas propuestas para la medición de contaminación lumínica mediante rastreo de satélites, son convenientes o aplicables al análisis de impacto sobre un territorio determinado?

- Palabras clave y sinónimos: se presenta el siguiente cuadro resumen con los términos a utilizar para la revisión

Palabra clave	Traducción al inglés
Rastreo de satélites	Satellite tracking
Brillo de satélite	Satellite Flare Satellite glint
Minería de datos	Data Mining
Regresión lineal	Linear Regression
Constelación de satélites	Satellite constellation
Observación astronómica	Astronomical observation
STARLINK	STARLINK
Aprendizaje supervisado	Supervised Learning

Tabla 1: Palabras a incluir en la revisión sistemática

- **Intervención:** En el contexto de la revisión sistemática se revisarán todas las propuestas y enfoques existentes sobre el impacto de la contaminación del brillo de los satélites en cualquier área de la tierra, para analizar y encontrar la que más se ajuste, tanto en resultados positivos como negativos.
- **Control:** Para la revisión se ha obtenido información de noticias y documentos recientes acerca del problema, popularizados por los lanzamientos de la red de satélites STARLINK de la empresa. SpaceX.
- **Resultados:** Los resultados esperados de esta revisión son detectar aquellos estudios que hayan cubierto el impacto de interferencia de los satélites en la observación astronómica para luego analizarlos y utilizar aquellos que implementen métodos o procedimientos afines a la presente investigación.
- **Medida de salida:** Para medir los resultados se iniciará por obtener el número de enfoques, algoritmos y metodologías que se puedan implementar o se puedan utilizar como referencia para la presente investigación.
- **Población:** La población está compuesta por las publicaciones de los repositorios seleccionados para la búsqueda con temas análogos.
- **Aplicación:** Los beneficiarios de dicha revisión son los analistas de datos.
- **Diseño experimental:** No se realizará el meta análisis en la revisión sistemática.

1.9.2 Selección de fuentes

1.9.2.1 Definición del criterio de selección de fuentes

Para la sección de las fuentes se utiliza el criterio del autor, quien, basándose en su experiencia profesional en análisis y presentación de datos, sugerirá la lista sobre las cuales se va a realizar la revisión. Otros criterios para tomar en cuenta es la accesibilidad de los documentos a través de la web y la utilización de motores de búsqueda que permitan consultas avanzadas.

1.9.2.2 Lenguajes de estudio

El lenguaje a utilizar para las búsquedas será el inglés. Sin embargo, el informe de la revisión se presenta en español.

1.9.2.3 Identificación de fuentes

- Método de selección de fuentes: Se seleccionan aquellas fuentes que puedan ser accedidos desde internet que cumplan con los criterios establecidos y que gocen con un rango de confiabilidad (primer cuartil) en la plataforma “Scimago Journal”. También se toman en cuenta los estudios respaldados por las universidades que se encuentren en el top 100 mundial, listadas en www.topuniversities.com.
- Lista de fuentes:
 - ✓ IEEE
 - ✓ Google Scholar.
 - ✓ Astrophysical Journal Letters
 - ✓ Astronomy and Astrophysics
 - ✓ American Astronomical Society
 - ✓ Microsoft Academic
 - ✓ ScienceDirect
 - ✓ IOP Science
- Cadenas de búsqueda: Combinaciones de los operadores lógicos “O” e “Y” sobre los términos seleccionados y sus sinónimos en inglés. Las cadenas a utilizar serán: (Satellite AND constellation OR Glint OR Flare), (STARLINK AND flare OR glint), (Satellite tracking AND astronomical observation), (Impact AND Satellite Constellation AND Light pollution), (Astronomical

observation AND flare OR glint), (Satellite orbit AND data mining), (Satellite tracking AND linear regression), (Satellite orbit AND supervised learning)

1.9.2.4 Selección de fuentes después de la evaluación

Se realiza un refinamiento de todas las fuentes que fueron seleccionadas y que cumplen con los criterios de selección. Se eliminarán los estudios que no sean provechosos para la investigación.

1.9.2.5 Selección de fuentes después de la evaluación

Todas las fuentes que cumplan con los criterios mencionados anteriormente serán aprobadas.

1.9.3 Selección de estudios

1.9.3.1 Procedimiento para selección de estudios

El proceso de selección de estudios se detalla en la Figura 3

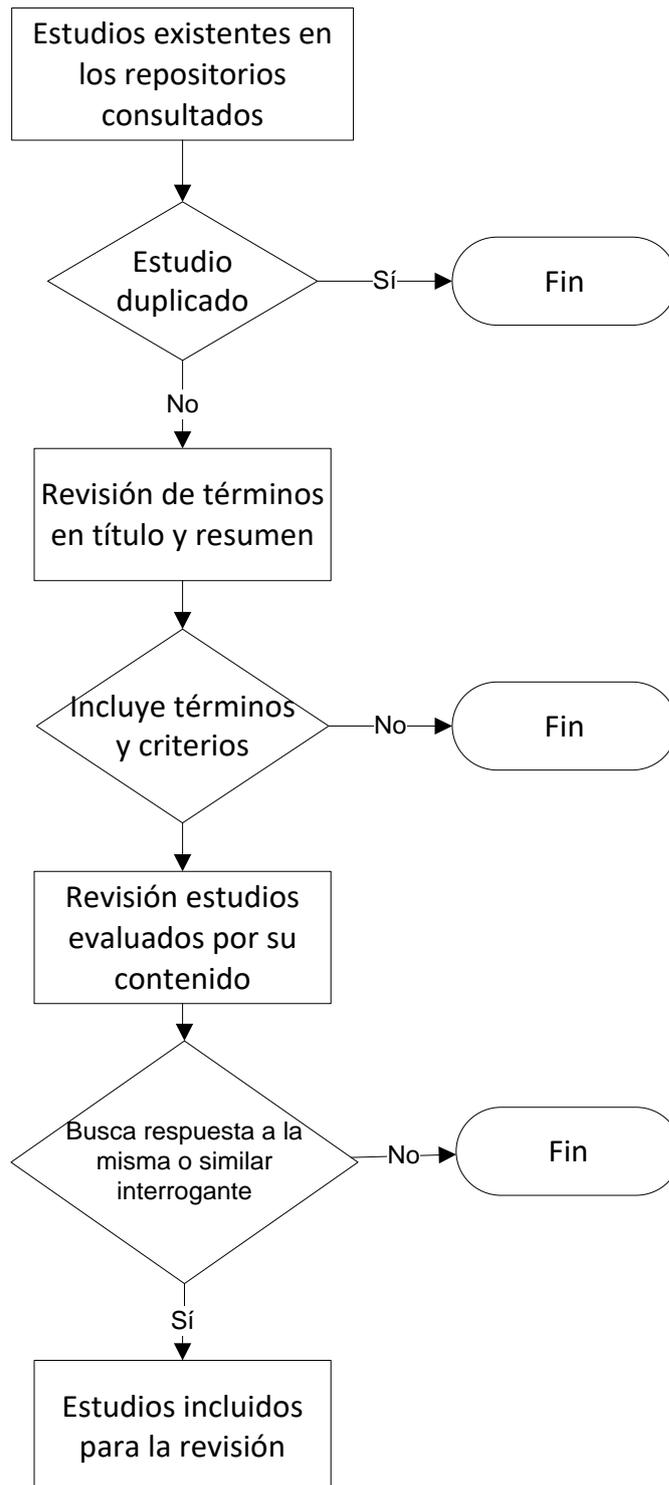


Figura 3: Diagrama del proceso de revisión de literatura. Fuente: Elaboración propia

1.9.3.2 Definición del criterio de inclusión y exclusión de estudios

En primera instancia se realiza una evaluación de las palabras clave sobre el título y resumen ejecutivo de los estudios. Esto permite excluir de la revisión aquellos que no tengan coincidencias con los términos empleados.

Luego se hace un análisis detallado del contenido del estudio y sus conclusiones, para así filtrar de manera más detallada si los criterios encontrados son realmente referentes al tema o si por el contrario no se utilizan en el mismo contexto.

1.9.3.3 Definición del tipo de estudio

Se utilizarán aquellos estudios que provengan de las fuentes citadas en la sección 1.9.2.3 y que cumplan con los criterios definidos en la tabla de palabras clave de la sección 1.9.1.2.

1.9.4 Extracción de la información

1.9.4.1 Definición del criterio de inclusión y exclusión de la información

La información que se considere relevante en los estudios es aquella que incluya técnicas, algoritmos, métodos y fórmulas que expliquen el cómo detectar o estimar las órbitas, impacto en observación, magnitud aparente o cómo evaluar la validez de estos métodos.

1.9.4.2 Formulario para la extracción de la información

En la Tabla 2 se observa el formato del formulario utilizado para documentar la extracción de la información. La primera parte consta de la identificación del estudio, donde se incluye el título, autor y año de publicación. En la segunda parte una descripción general sobre el estudio, además de las conclusiones. Finalmente, en la tercera parte se documentan los aspectos que el investigador de este trabajo determina importante destacar del estudio.

Identificación	Título
	Año
	Autores
Contenido	Resumen del contenido y conclusiones del estudio
Aspectos importantes	

Tabla 2: Formulario de extracción de información. Fuente: Elaboración propia.

1.9.4.3 Extracción de resultados objetivos y subjetivos

En esta sección se exponen los estudios encontrados utilizando la tabla mostrada anteriormente. Además, luego se hace un análisis

Identificación	
Título	The Low Earth Orbit Satellite Population and Impacts of the SpaceX Starlink Constellation
Año	2020
Autores	McDowell J.
Contenido	
<p>El estudio hace una estimación del impacto de los satélites Starlink de la empresa SpaceX en la observación astronómica profesional. Hace un recuento del total de objetos en el cielo dividiendo estos satélites en tres categorías de acuerdo a la altura de la órbita y analiza la información de rastreo de todos ellos.</p> <p>El trabajo realiza una suposición de la distribución de los aproximadamente 12000 satélites a ser desplegados por SpaceX en total.</p> <p>Con esto determina que:</p> <p>Para latitudes bajas los observatorios no se verían tan afectados, como aquellos que estén en latitudes más altas (45-55 grados).</p> <p>Se espera que en el futuro cercano pueda haber una densidad de alrededor de 0.01 objetos por grado cuadrado (deg^2) en el cenit del observador.</p>	

<p>Las observaciones más afectadas son aquellas donde se necesiten mayores tiempos de exposición, que se deban realizar al atardecer o amanecer, observatorios de latitud más alta y en ciertas épocas del año.</p>	
<p>Aspectos importantes</p>	
<p>El estudio hace utiliza métodos y fórmulas de cálculo que fundamentan las conclusiones del estudio, así como información relevante de los objetos que orbitan la tierra.</p> <p>Este estudio es muy completo en cuestión a los satélites de SpaceX. Sin embargo, no ahonda en profundidad la afectación de los satélites ya existentes antes del programa Starlink.</p>	
<p>Extracción de resultados objetivos</p>	
<p>Metodología del estudio</p>	<p>El autor hace una revisión de los objetos que existen en registrados en catálogo estadounidense de objetos espaciales y realiza una estimación de los satélites por desplegar en las misiones futuras.</p>
<p>Resultados del estudio</p>	<p>Determina que las observaciones más afectadas son las de alta exposición y los observatorios en latitudes más altas.</p>
<p>Problemas del estudio</p>	<p>No se encuentra ninguno</p>
<p>Extracción de resultados subjetivos</p>	
<p>Información a través de los autores</p>	<p>No se contactó al autor del trabajo.</p>
<p>Impresiones generales</p>	<p>El trabajo es suficientemente específico en cuanto a demostrar el impacto de los satélites en la observación profesional.</p>

Tabla 3: Primera fuente literaria investigativa. Información obtenida de McDowell (2020)

Identificación	
Título	Mega-Constellations of Satellites and Optical Astronomy
Año	2020
Autores	Seitzer, P.
Contenido	
<p>Este trabajo calcula el impacto que tuvieron los primeros lotes de satélites Starlink. Indica que estos últimos son más brillantes que el otro 99% de objetos que se encuentran actualmente en órbita. Algunos de los impactos que indica en su estudio son: La pérdida de información en pixeles, interferencia en aparatos electrónicos, imágenes fantasmas y posibles imágenes residuales.</p>	
Aspectos importantes	
<p>Este estudio realiza cálculos preliminares del problema, utiliza datos de los objetos en el catálogo público de objetos espaciales y la información públicamente conocida tras los lanzamientos de SpaceX a finales de 2019 e inicios de 2020.</p>	
Extracción de resultados objetivos	
Metodología del estudio	<p>El autor utiliza datos de observaciones en el observatorio de Cerro Tololo en Chile y la información pública de la ubicación de los satélites para identificar la cantidad de satélites que serán visibles por los astrónomos haciendo énfasis en la ubicación del observatorio mencionado.</p>
Resultados del estudio	<p>Se concluye que los satélites en órbitas bajas solamente tendrían un impacto en los periodos de amanecer y anochecer.</p> <p>El autor también expresa que debe existir preocupación por la situación ya que se espera que la cantidad de satélites continúe en crecimiento.</p>

Problemas del estudio	El estudio está realizado en formato de presentación por lo que presenta algunas ideas que no se describen detallada mente correctamente en el documento publicado.
Extracción de resultados subjetivos	
Información a través de los autores	No fue solicitado.
Impresiones generales	Este estudio sirve como punto de partida para análisis más detallados.

Tabla 4: Segunda fuente literaria investigativa. Información obtenida de Seitzer (2020)

Identificación	
Título	Impact of satellite constellations on astronomical observations with ESO telescopes in the visible and infrared domains.
Año	2020
Autores	Hainaut, O., Williams, A.
Contenido	
Esta investigación realiza un análisis y estimación del posible impacto que tendrían los satélites en las observaciones astronómicas terrestres en el futuro basándose en los datos recolectados desde diversas fuentes sobre los lanzamientos de constelaciones de satélites que se tienen previstas, específicamente 18 constelaciones en desarrollo por SpaceX, Amazon, OneWeb entre otras, cuya cantidad conjunta de satélites es superior a 26 mil.	
Aspectos importantes	
Esta investigación asume que los satélites tendrán una distribución uniforme sobre la tierra, además su magnitud es estimada utilizando un modelo simple basado en observaciones reales para obtener una estimación conservadora de los efectos.	
Extracción de resultados objetivos	
Metodología del estudio	Mediante esta investigación se pretende cuantificar el impacto de grandes constelaciones de satélites en la

	observación astronómica del ESO (European Southern Observatory), por medio de estimaciones basadas en observaciones actuales y considerando el aumento de la cantidad de satélites según la planificación de lanzamientos que próximamente se efectuarán.
Resultados del estudio	A partir de este estudio se determina el impacto del paso de satélites en las observaciones realizadas con distintos tipos de telescopios y dependiendo también del tiempo de exposición. Para observaciones de corta duración y con telescopios pequeños no se verán afectadas de forma significativa. Observaciones de mediana duración se afectarán hasta un 0.5% en los atardeceres. Las observaciones de larga exposición al inicio o al final de la noche se afectarán hasta en un 3%. Las observaciones de muy amplio campo con grandes telescopios se verán afectadas hasta en un 30%.
Problemas del estudio	El análisis se basa en estimaciones y se simplifican algunas variables como la distribución real de los satélites
Extracción de resultados subjetivos	
Información a través de los autores	No fue solicitado.
Impresiones generales	A partir de las medidas de mitigación recomendadas por el autor se destaca la necesidad de tener disponible los datos actualizados y precisos de los elementos orbitales de todos los satélites para poder predecir la posición exacta y horarios de los satélites

Tabla 5: Tercera fuente literaria investigativa. Información obtenida de Hainaut (2020)

Identificación	
Título	Concerns about ground based astronomical observations: A step to Safeguard the Astronomical Sky

Año	2020
Autores	Gallozi, S., Scardia, M., Maris, M.
Contenido	
<p>Tiene como objetivo destacar el impacto de las observaciones astronómicas terrestres en diferentes ventanas del espectro electromagnético provenientes del despliegue de flotas de satélites de telecomunicaciones. Se presta especial atención al problema del hacinamiento del espacio circunterrestre por objetos en órbita de tamaño medio o pequeño.</p>	
Aspectos importantes	
<p>El estudio está más enfocado a explicar la importancia de un cielo limpio y defender el derecho de la humanidad a vivir en un entorno sin contaminación. En una de sus conclusiones indica que todos estos lanzamientos deberían ser aplazados hasta que se tenga una mejor regulación en el ámbito legal, para proteger la inversión en la observación en tierra y los mismos derechos de los seres humanos de vivir en condiciones óptimas para la observación y estudio del cielo.</p>	
Extracción de resultados objetivos	
Metodología del estudio	<p>El estudio realiza un conteo de cuántos satélites existen y teoriza la cantidad que van a ser lanzados en el futuro cercano. Clasifica los satélites de acuerdo al ángulo de reflexión que podrían tener y calcula el grado de interferencia que pueden llegar a tener dadas las especificaciones de los fabricantes.</p> <p>Hace un análisis detallado de los efectos negativos que los satélites tienen en las diferentes áreas directamente impactadas por el despliegue de estos efectos, explicando el qué y cómo producen una obstrucción a los estudios actuales.</p>
Resultados del estudio	<p>Llega a la conclusión que el peor de los casos se experimentará durante el verano, más allá de la mitad del cielo, en latitudes más altas, donde los satélites serán visibles durante toda la noche.</p>

	El mejor escenario será durante el invierno, a media noche cuando el cielo estaría prácticamente libre de satélites a excepción del horizonte.
Problemas del estudio	No se encuentran problemas
Extracción de resultados subjetivos	
Información a través de los autores	No fue solicitado.
Impresiones generales	El estudio hace una parada por cada una de las áreas afectadas directamente por el despliegue de nuevos satélites. También es una carta abierta para alentar a otros a unirse contra la contaminación que estos objetos en el cielo puedan causar.

Tabla 6: Cuarta fuente literaria investigativa. Información obtenida de Gallozi (2020)

Identificación	
Título	Research on Satellite Orbit Prediction Based on Neural Network Algorithm
Año	2019
Autores	Ren, H., Chen, X., Guan, B., Wang, Y., Liu, T., Peng, K.
Contenido	
<p>Esta investigación se resalta la importancia de una predicción lo más exacta posible de las órbitas de los satélites para prever y evitar posibles colisiones. El estudio propone una alternativa a la predicción tradicional mecánica, en la cual utilizan la técnica de minería de datos llamada redes neuronales, la cual consiste en el entrenamiento del modelo con datos históricos de TLE obtenidos de NORAD, Mando Norteamericano de Defensa Aérea.</p>	
Aspectos importantes	
<p>Esta investigación utilizó datos experimentales históricos del sistema de satélites IRIDIUM compuesto por 66 satélites activos. Dichos datos históricos comprenden un rango de años desde 1982 hasta 2004. Además, se trabajó con el algoritmo de</p>	

minería de datos Red Neuronal LSTM, la cual es un tipo de red neuronal recurrente (RNN).	
Extracción de resultados objetivos	
Metodología del estudio	Mediante esta investigación se pretende definir un modelo de predicción de órbitas de satélites alternativo al método tradicional, con el cual se logró un nivel de exactitud mayor utilizando la técnica de minería de datos llamada Redes Neuronales, la cual utiliza datos históricos para aprender patrones y realizar las predicciones de las variables estudiadas.
Resultados del estudio	A partir de este estudio se determina que el modelo de redes neuronales tiene un margen de error aceptable y se considera que se puede llegar a mejorar.
Problemas del estudio	El estudio demuestra que el margen de error es incremental en el tiempo, por lo que los resultados son más confiables en los primeros días y menos confiable en predicciones a mediano plazo.
Extracción de resultados subjetivos	
Información a través de los autores	No fue solicitado.
Impresiones generales	Este estudio da una alternativa en la predicción de órbitas satelitales utilizando modelos de minería de datos, con la cual se pueden tener datos relativamente confiables en el corto plazo, lo que puede resultar útil en la investigación para la predicción de la ubicación de los satélites y su impacto en las observaciones astronómicas.

Tabla 7: Quinta fuente literaria investigativa. Información obtenida de Ren (2019)

Identificación	
Título	Mining Two-line-element data to detect orbital maneuver for

	satellite
Año	2019
Autores	Bai, X., Liao, C., Pan, X., Xu, M.
Contenido	
<p>Esta investigación plantea el análisis de los datos TLE (two-line element) con el fin de detectar maniobras de los satélites para mantenerse en las órbitas deseadas y así llevar a cabo su función. Para la detección de las maniobras realiza un estudio analizando datos históricos de tres satélites, los cuales son: “YAOGAN-9”, “TIANHUI-1” y “Envisat” los cuales son fácilmente detectables y proveen la información necesaria para validar el modelo.</p>	
Aspectos importantes	
<p>Este tipo de análisis busca identificar de mejor manera los micro movimientos de los satélites. Estos son muy difíciles de detectar mediante métodos de vigilancia tradicionales y que afectan la predicción de las órbitas de los mismos.</p>	
Extracción de resultados objetivos	
Metodología del estudio	<p>El artículo utiliza tres métodos de agrupamiento de minería de datos. El K-Means, clusters jerárquicos y Fuzzy C-Means. Estos métodos son utilizados para descubrir patrones en los datos y permitan hacer predicciones y se utilizan cuando no existe una clasificación clara en los datos y por consiguiente, deben ser divididos en primera instancia.</p>
Resultados del estudio	<p>Se determinó que el método K-Means es el mejor para clasificar las maniobras medias y grandes en escala (con un 1% de error en la clasificación). Mientras que Fuzz-C Means es la mejor técnica para detectar las maniobras pequeñas de manera más eficiente (con un 4% de error).</p>
Problemas del estudio	<p>Solo utilizan datos de 3 satélites en específico, por lo que los resultados podrían no ajustarse a otros satélites.</p>
Extracción de resultados subjetivos	
Información a través	No fue solicitado.

de los autores	
Impresiones generales	El artículo provee una metodología adicional para mejorar la predicción de las órbitas de los satélites mediante la detección de maniobras ejecutadas automáticamente por los satélites. Utilizar esta metodología aumenta la fiabilidad de las predicciones en las órbitas. Sin embargo, sería necesaria la aplicación de varias técnicas con el fin de predecir las maniobras con el menor error posible.

Tabla 8: Sexta fuente literaria investigativa. Información obtenida de Bai (2019)

1.9.5 Análisis de resultados

1.9.5.1 Resultados de cálculo estadístico

No se realizaron cálculos estadísticos.

1.9.5.2 Presentación de resultados

La siguiente tabla de resultados muestra los tipos de estudios seleccionados que hayan realizado mediciones del impacto del tránsito satelital en la órbita terrestre.

	Brillo de satélites	Starlink	Minería de datos	Total
No. estudios seleccionados	1	3	2	6

Tabla 9: Presentación de resultados de revisión sistemática

Como criterio de extracción de los resultados y la información de los mismos se utiliza como criterio si el estudio tiene menos de 5 años desde su publicación.

1.9.5.3 Análisis de sensibilidad

No fue aplicado

1.9.5.4 Gráficos

No fue aplicado

1.9.5.5 Comentarios finales

- Número de estudios: 9 estudios encontrados, 5 seleccionados.
- Sesgo de búsqueda, selección y extracción: Todos los estudios que fueron publicados antes del año 2015 han sido descartados.
- Sesgo de publicación: No fue definido.
- Variación entre revisores: No hay variación.
- Aplicación de resultados: Los estudios seleccionados comparten de manera general la misma metodología de estudio. Esto permite la combinación de estas técnicas para determinar el impacto de los satélites en órbita en la observación astronómica y comparar los resultados para la creación de un mejor modelo.
- Recomendaciones: Ninguna.

Capítulo 2. Marco Teórico o Conceptual

En este capítulo se presenta el apoyo conceptual generalizado basado en la revisión de literatura realizada en el capítulo el estado de la cuestión. Estos estudios son tomados en consideración en la elaboración de esta investigación sobre el análisis de la órbita de los satélites para la estimación del impacto de estos en la observación astronómica.

2.1 Mapa conceptual

bright clustering constellation contour deg elevation exposure
 flares ground horizon k-means leo lstm magnitude maneuver mining model network
 neural observations orbital satellites
 starlink tle zenithal

Figura 4: Nube de conceptos. Fuente: Elaboración propia



Figura 5: Mapa conceptual. Fuente: Elaboración propia

A continuación, se detallan los conceptos más relevantes para la investigación, que se encontraron, los mismos se describirán siguiendo la jerarquía de lo general a lo específico.

2.2 Observación astronómica

Según la RAE la observación es la acción examinar atentamente y astronomía es la ciencia que trata de los astros, de su movimiento y de las leyes que lo rigen.

La observación astronómica es una rama de la astronomía que recopila y almacena información sobre el universo observable. Cheng (2009) explica que esta conlleva el estudio de los cuerpos celestes por medio del uso de telescopios y otras herramientas que permiten el estudio del espectro electromagnético. Este espectro varía de acuerdo a la longitud de onda de la luz abarcando desde las ondas de radio, radiación infrarroja, rayos ultravioletas, rayos x, rayos gamma y el espectro de luz visible.

Para efectos de este trabajo se hará referencia al término de observación astronómica a la técnica de obtención de datos mediante el uso de telescopios ópticos, los cuales son aquellos que capturan ondas de luz visible.

2.2.1 Tiempo de exposición

El tiempo de exposición es la cantidad de tiempo en la que la cámara captura luz desde la fuente (Chen et al., 2020).

En astronomía este tiempo puede variar desde algunos segundos hasta varias horas como lo indican Hainaut y Williams (2020). Y se realiza con el fin de obtener la mayor cantidad de detalle posible de un astro celeste. Este tiempo varía de acuerdo a la observación y el instrumento realizado.

Cuanto mayor es el tiempo de exposición, mayor es el riesgo que una imagen sea perjudicada por el tránsito de un satélite. Los telescopios más grandes se suelen utilizar con mayores tiempos de exposición y por lo tanto son los más afectados dentro del ámbito de esta investigación, como lo hace ver Hainaut (2020) en su estudio: "... las exposiciones de imágenes ultra anchas en un telescopio muy grande, se verían significativamente afectadas, con un 30-40% de dichas exposiciones comprometidas durante las primeras y últimas horas de la noche"

2.2.2 Azimut

Este es el “ángulo que con el meridiano forma el círculo vertical que pasa por un punto de la esfera celeste o del globo terráqueo” RAE (2020).

En otras palabras, es la ubicación de un objeto celeste, medido en dirección a las manecillas del reloj desde el norte del observador y se mide en grados. De esta manera el azimut norte equivaldría a 0° , este 90° , sur 180° y oeste 270° .

Por lo general, para ubicar un astro desde la superficie de la tierra, el azimut se utiliza en conjunto con la elevación (también medida en grados) donde estas coordenadas forman un ángulo que el observador toma como punto de referencia (Ver Figura 6).

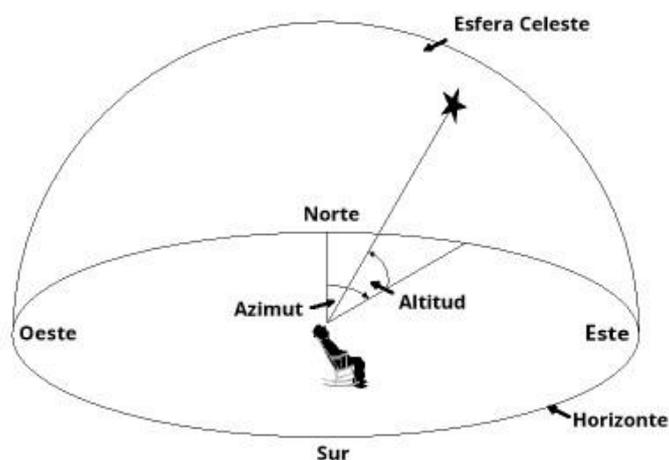


Figura 6: Ubicación de un objeto utilizando el azimut. Adaptada al español del sitio web heavens-above.com

2.2.3 Cénit

La RAE define cénit en el ámbito astronómico como: “Intersección de la vertical de un lugar con la esfera celeste, por encima de la cabeza del observador.”

En tanto se puede entender como un punto imaginario que apunta hacia arriba de una determinada ubicación opuesta a la fuerza gravitacional. Por lo que se refiere a lo que está directamente sobre el observador.

El cénit se utiliza en astronomía al igual que el azimut y la elevación para calcular la posición de un objeto en el cielo. Sin embargo, la diferencia entre el cénit

y la elevación radica en el origen del cálculo del ángulo, dado que para el primero se toma como punto de partida la vertical, como se muestra en la Figura 7.

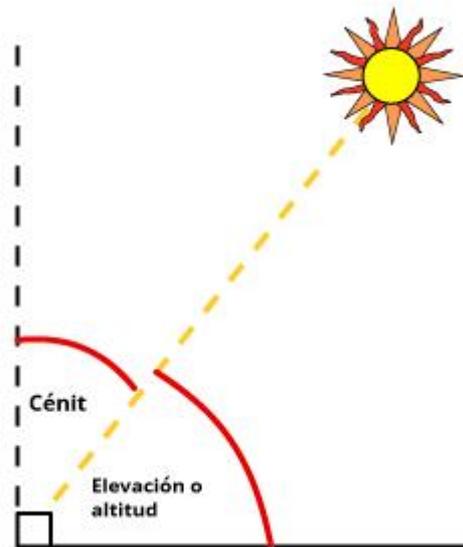


Figura 7: Ilustración del cénit. Adaptada al español de <https://www.pveducation.org/pvcdrom/properties-of-sunlight/elevation-angle>

2.2.4 Constelación de satélites

Una constelación de satélites se describe como: “un número de satélites similares, de un tipo y función similares, diseñados para ser similares, complementarios, que orbitan por un propósito compartido, bajo un control compartido” (Wood, 2003)

Estas constelaciones, según la unión astronómica internacional (IAU), han sido utilizadas para la navegación y geodesia (estudio de la forma de la tierra). Las más aplicaciones comunes son: las redes GPS, telefonía satelital y la observación de la tierra en general (meteorológica, relieve, entre otras).

Recientemente varias compañías planean desplegar constelaciones de satélites de larga escala con el fin de proveer internet satelital que conecte las máquinas y sistemas de manera directa. La IAU explica que la ventaja para la existencia de dichas constelaciones es que debido a la cercanía entre los satélites y su baja órbita se puede reducir el tiempo que tardan las señales en viajar de un punto a otro, propiciando una comunicación más eficiente.

2.2.4 Magnitud aparente

“La magnitud aparente describe el brillo de un objeto visto o capturado en imagen desde la tierra. La magnitud aparente de un objeto se determina por ambas, la cantidad de luz que el objeto emite y la distancia desde el observador. Un objeto con una baja magnitud aparente se ve brillante dado que emite mucha luz, está cerca o ambos. Un objeto distante que emite una cantidad masiva de luz puede tener una magnitud aparente más baja (más brillante) que otros que están mucho más cerca pero que emiten menos luz” (Thompson & Thompson, 2005, P. 88).

La magnitud se indica con una “m” minúscula. Así por ejemplo una estrella con una magnitud aparente de 4.7 se registra como “4.7m” durante las observaciones.

Según la asociación americana de observadores de estrellas variables (AAVSO, por sus siglas en inglés). Se ha estimado que el ojo humano percibe cambios en 1 unidad de magnitud como 2.5 veces más brillante. Por lo tanto, un cambio de 5 unidades de magnitud sería el equivalente a 2.5^5 o aproximadamente 100 veces más brillante. Por ende, se ha definido que un cambio en 5 niveles de magnitud corresponde al factor de 100 en magnitud aparente. En esta escala algunos objetos son tan brillantes que tienen magnitudes negativas.

En la Figura 8 se muestran algunos ejemplos comunes en una escala, que resultan familiares al lector.

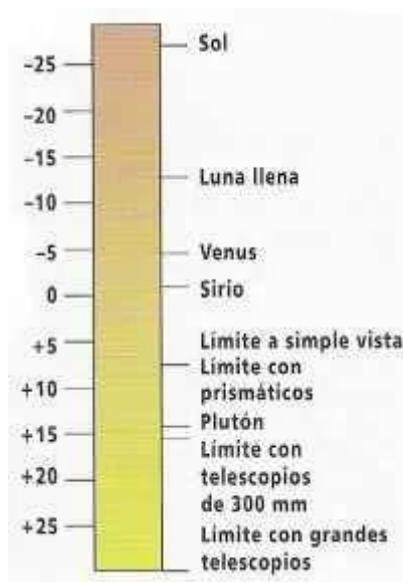


Figura 8: Ejemplos de magnitud aparente. Tomada de <https://www.astronomia-iniciacion.com/magnitudes.html>

Para determinar la magnitud aparente de un objeto, este debe observarse en las condiciones y momentos dados. Sin embargo, esta metodología no es posible en el alcance de la presente investigación.

Por lo tanto, para poder estimar la magnitud visual de un satélite de manera aproximada, se suele utilizar una metodología que asume que el satélite es una esfera difusa.

Para determinar el comportamiento fotométrico de una esfera difusa y por lo tanto de la magnitud, Hejduk (2011) propone la utilización de dos funciones de fase denominadas “especular-difuso”.

$$F(\varphi)_1 = \frac{2}{3\pi^2} [(\pi - \varphi) \cos \varphi + \sin \varphi]$$

$$F_2(\varphi) = \frac{1}{4\pi}$$

Estas dos fórmulas son utilizadas para el cálculo final.

$$M_v = -26.74 - 2.5 \log_{10}(A\rho[\beta F_1(\varphi) + (1 - \beta)F_2(\varphi)]) + 5 \log_{10}(R)$$

En donde -26.74 es la magnitud aparente del sol. F_1 es la función de difusión de fase. F_2 es la función especular fase de la esfera. M_v es el brillo resultante. φ es el ángulo de fase (en radianes). A es el área de sección transversal. ρ es el albedo. R es la distancia hasta el objeto. En este contexto $\beta = 1$ representaría una esfera puramente difusa y $\beta = 0$ una esfera especular. Sin embargo, dado que los satélites no suelen ser totalmente difusos o especulares, se combinan los coeficientes.

2.3 Rastreo satelital

Desde el primer lanzamiento orbital en 1957, el número de objetos y escombros orbitando la tierra ha ido en incremento. La NASA en el volumen 24 de su boletín trimestral de noticias, indica que existen en órbita aproximadamente veinte mil objetos y, de estos, alrededor de cinco mil son satélites y/o vehículos espaciales.

Estos objetos son en principio monitoreados ya que son una amenaza para otros satélites y vehículos activos, dada la alta velocidad a la que orbitan, poniendo en peligro servicios cruciales para la humanidad (Virgili et al., 2016).

Por lo anterior descrito, la fuerza aérea de los estados unidos mediante su red de vigilancia espacial, mantiene un catálogo e información actualizada sobre los objetos en órbita descubiertos que tengan más de 10 centímetros de diámetro. Esta base de datos es pública y está disponible a través del sitio web space-track.org. Utilizando esta base de datos es posible para los usuarios registrados obtener información de localización y utilizarlos para análisis diversos.

Dicha base de datos será la utilizada en la realización del presente trabajo. Estos datos son la base del análisis y estimaciones a realizar.

2.3.1 Órbita baja terrestre (LEO)

Según McDowell (2020) La órbita baja terrestre o Low Earth Orbit en inglés, es generalmente considerada desde los 80 y 100 km (más bajo de ese punto no pueden permanecer en órbita) y hasta una altitud de hasta los 2000 km de altitud donde la intensidad de los cinturones de radiación hace más difícil que los satélites puedan operar.

Adicionalmente McDowell (2020) también indica que estos satélites, bajo buenas condiciones atmosféricas, pueden llegar a ser visibles a simple vista. Esto es lo que lleva a los astrónomos a preocuparse por el futuro de la observación espacial.

2.3.2 Reflejo satelital

Los reflejos satelitales son causados cuando la luz del sol choca con un satélite y este refleja la luz del sol hacia la tierra. (Schaefer et al., 1987). Estos reflejos pueden darse debido la superficie reflectante del satélite, como por ejemplo los paneles solares o las antenas.

“... puede definirse como un breve resplandor de luz que dura de 1 a 15 s. Sin embargo, incluso esta definición es algo arbitraria, ya que no hay una fórmula consistente para cada reflejo. Sí se sabe cuándo se puede esperar, incluso se puede ver desde el interior de una casa, entre árboles, en una noche en la que el

cielo está completamente cubierto por una capa delgada y alta, o incluso en la luz del día” (Maley & Pizzicaroli, 2003)

Estos reflejos solamente ocurren cuando hay una alineación determinada para que la luz del sol llegue al observador. Es por esto que la posición del satélite, la hora del día y la época del año (debido a la inclinación en el eje de la tierra) tienen un impacto en el reflejo obtenido. En la Figura 9 se muestra un ejemplo de los reflejos satelitales y cómo se producen.

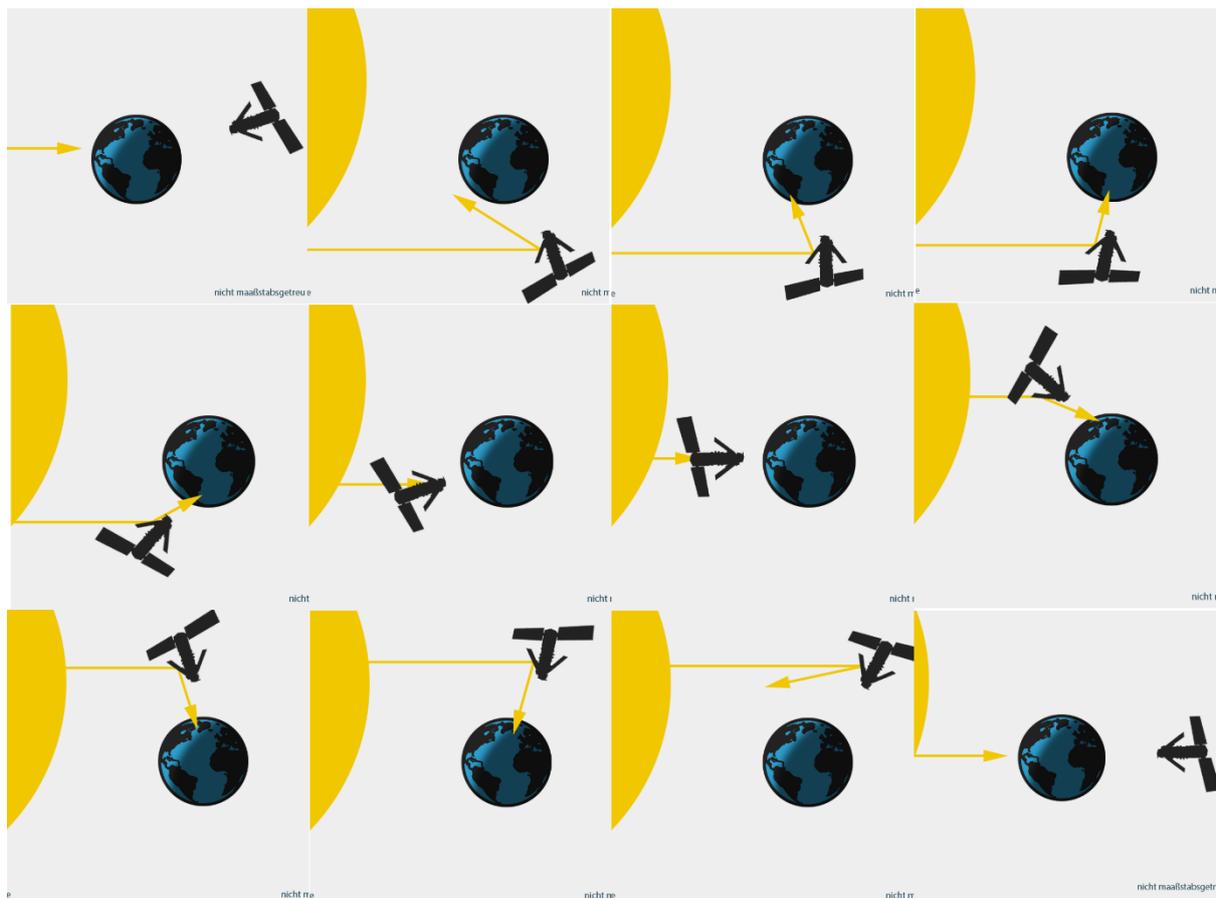


Figura 9: Visualización del reflejo satelital Tomada de <https://www.high-iso.de/blog/tutorials/iridium-flares-fotografieren/>. Video convertido a imágenes para una mejor interpretación.

2.3.3 Elementos de dos líneas (TLE)

Los elementos de dos líneas (two-line elements) han sido aceptados como un estándar para la presentación de datos orbitales, sin embargo, se muestran de una manera no intuitiva para la caracterización del tipo y forma de la órbita. Esta información es generada por el Comando Norteamericano de Defensa Aeroespacial

(NORAD, por sus siglas en inglés) y se utiliza como fuente para determinar y predecir las órbitas de los objetos que orbitan la tierra (Carvalho, Estela, & Langer 2020)

En la Figura 10 se presenta un ejemplo de un elemento de dos líneas.

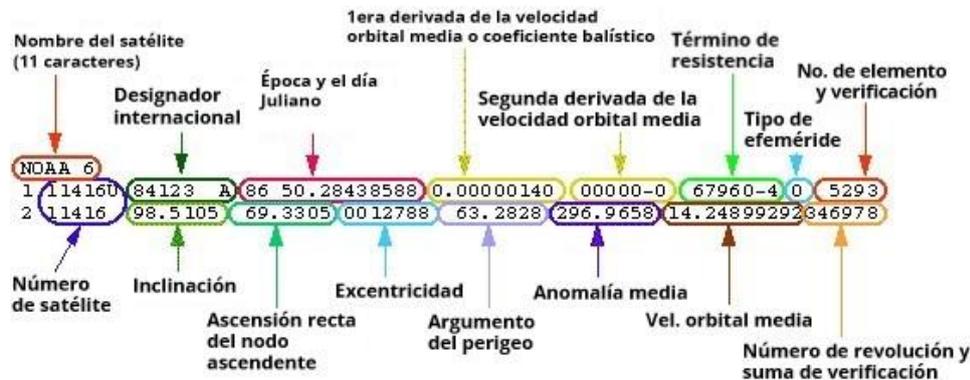


Figura 10: Ejemplo de un elemento de dos líneas. Adaptada al español de https://spaceflight.nasa.gov/realdata/sightings/SSApplications/Post/JavaSSOP/SSOP_Help/tle_def.html

A continuación, se detalla el significado de cada uno de los elementos mostrados en la imagen anterior. Todos los conceptos han sido traducidos al español directamente de las definiciones proveídas por la NASA.

Nombre del satélite

(NOAA 6) Este es simplemente el nombre asociado con el satélite.

Designador internacional

(84 123A) Primeros 2 caracteres representan el año del lanzamiento, siguientes 3 son el número (123) de lanzamientos para ese año y el último carácter (A) identifica el número de componente del lanzamiento (A=1)

Época y el día Juliano

El día Juliano es el número de días que han pasado en un año en particular. Por ejemplo, la fecha arriba muestra "86" como el año 1986 y 50.28438588 el cual es una fracción indicando que han pasado un poco más de 50 días después del 1 de enero de 1986. Esto daría la siguiente fecha resultante

1986/050:06:49:30.94 siguiendo el siguiente formato:

Año/Número de días:Horas:Minutos:Segundos.

El cálculo es el siguiente:

Se comienza con 50.28438588 días

Días = 50

50.28438588 días - 50 = 0.28438588 días

0.28438588 días x 24 horas/día = 6.8253 horas (por lo tanto, horas = 6)

6.8253 horas - 6 = 0.8253 horas

0.8253 horas x 60 minutos/hora = 49.5157 minutos (por lo tanto, minutos = 49)

49.5157 - 49 = 0.5157 minutos

0.5157 minutos x 60 segundos/minuto = 30.94 segundos (Segundos = 30.94)

Coeficiente balístico

(0.00000140) También llamado "Primera derivada de la velocidad orbital media". Es la tasa diaria de cambio en el número de revoluciones que un objeto completa cada día dividido entre 2. La unidad es revoluciones/día.

Segunda derivada de la velocidad orbital media

(00000-0 = 0.00000) La segunda derivada del movimiento medio es un término de arrastre utilizado para modelar la disminución de la órbita terminal. Mide la segunda derivada del movimiento medio diario, dividido entre 6. La unidad de medida es revoluciones/día³. Los últimos 2 caracteres indican si después del número el valor debe ser aplicado a la potencia de 10.

Entonces por ejemplo el valor 12345-5 = 0.0000012345

Término de resistencia

(67960-4 = 0.000067960) También llamado el coeficiente de presión de radiación. La unidad de medida es R_e^{-1} donde R_e ese el radio de la tierra.

Los últimos 2 caracteres indican si después del número el valor debe ser aplicado a la potencia de 10.

Número de elemento y suma de verificación

(5293) Es un conteo de todos los TLEs generados por USSPACECOM para este objeto (en este ejemplo 529). El contador eventualmente llega a 999 lo cual hace que vuelva a iniciar en 1. El último dígito de la línea es el dígito de verificación de la línea 1 utilizado por los sistemas para verificar la integridad de la línea como tal.

(Suma de todas las cifras numéricas individualmente, añadiendo un 1 por cada signo negativo, y generando el módulo 10 de la cifra).

Tipo de efeméride

Este campo es para uso interno de SpaceTrack y denota el tipo de algoritmo utilizado en la generación del elemento de 2 líneas.

Número de satélite

(11416) Este es el número en el catálogo que USSPACECOM ha designado para el objeto. Una "U" indica que es un objeto desclasificado.

Inclinación (grados)

El ángulo entre el ecuador y el plano de la órbita.

Ascensión recta del nodo ascendente (grados)

Este es ángulo entre el equinoccio vernal y el punto donde la órbita cruza el plano ecuatorial (dirigiéndose al norte).

Excentricidad

(0012788) Una constante que define la forma de la órbita (0=circular, menor que 1 = elíptica). El valor proveído es la excentricidad media. Al valor debe agregársele un decimal 0, dicho de otra forma, restar el número a 0.

Argumento del perigeo (grados)

Es el ángulo entre el nodo ascendente y el punto de aproximación más cercano a la Tierra de la órbita (perigeo).

Anomalía media (grados)

El ángulo, medido desde el perigeo a la ubicación del satélite en la órbita referenciada a una órbita circular.

Velocidad orbital media (revoluciones/día)

(14.24899292) Este indica el promedio en el número de órbitas por día del objeto.

Número de revolución y suma de verificación

El número de órbitas en la fecha especificada. El último dígito es la suma de verificación, se calcula igual que el de la línea número 1.

Los elementos de dos líneas serán utilizados como base en el presente trabajo para la predicción de las órbitas de los satélites. Estos datos se pueden acceder públicamente desde el sitio space-track.org.

2.3.4 Maniobrabilidad satelital

Los satélites dependen principalmente de la radiación electromagnética reflejada o emitida desde la Tierra para obtener información acerca de la superficie de la Tierra o la atmósfera suprayacente. Hay ciertas maniobras orbitales que mantienen a los satélites en una órbita preestablecida u objetiva para que puedan completar sus tareas. (Bai, et al 2019)

Las maniobras en los satélites se implementan desde su despliegue y continúa en la fase de mantenimiento, con el fin de sostener el satélite en la órbita preestablecida para el fin deseado. Hay factores que perturban la órbita del satélite durante toda su vida. Sin embargo, para contrarrestar esto, se aplican cambios de velocidad al satélite. (Bai, et al 2019).

La Organización Europea para la Explotación de Satélites Meteorológicos (EUMESAT, por sus siglas en inglés) explica algunas las causas más comunes por las cuales la órbita de un satélite debe ser corregida. Estas son: ligeras asimetrías en el campo gravitacional de la Tierra debido al hecho de que la Tierra no es completamente esféricamente simétrica; la atracción gravitacional del sol y la luna; la presión de radiación solar; y, para los satélites en órbita terrestre baja, la resistencia atmosférica.

También expone que para la corrección del curso satelital se realiza una maniobra encendiendo propulsores incorporados en el satélite para cambiar la dirección o la velocidad del satélite. El combustible necesario para maniobrar está estimado de acuerdo a la vida útil del satélite, el cual al final de esta, hace una última maniobra para devolverlo a la tierra, reduciendo así la permanencia del satélite muerto en la órbita terrestre.

La maniobrabilidad es un factor adicional que añade complejidad a la tarea de predicción de órbitas. La determinación de dichas maniobras o el cálculo del impacto que tienen en el modelo a implementar no está incluida en el alcance de la

presente investigación. Por lo que en el modelo estos parámetros no serán tomados en cuenta, como se expondrá en las secciones siguientes.

2.3.5 Modelo SGP4/SDP4

Según Dong & Chang-Yin (2010) el modelo SGP4/SDP4 es un tipo de modelo de predicción de órbita desarrollado por NORAD (Comando de Defensa Aérea de América del Norte). En colaboración con la SSN (Red de Vigilancia Espacial de Estados Unidos). Estas entidades han generado el catálogo de objetos espaciales más grande en el mundo, el cual se comparte en el formato TLE (elemento de dos líneas) descrito anteriormente. Desde 1980 este modelo matemático y el correspondiente código creado en FORTRAN fueron publicados por el departamento de defensa de los Estados Unidos. Este código es constantemente revisado y actualizado.

Este modelo es uno de los modelos de perturbación general el cual utiliza algunas constantes físicas predefinidas para la rápida evaluación de los datos. Tener esta velocidad de procesamiento impacta la exactitud de las medidas ya que se evitan fórmulas y evaluaciones que pueden llegar a ser muy complejas, por lo que solamente las fuerzas más importantes son tomadas en cuenta (San-Juan, Pérez, San-Martín, & Vergara, 2017).

Los investigadores Xiao-Li & Yong-Qing (2019) realizaron varias mediciones y evaluaciones del algoritmo SGP4, luego de predecir las órbitas de 3953 satélites. Dividieron estos en cuatro tipos de órbitas: órbita baja (menor a 2000 km de altitud), órbita media (entre 2000 km y 36000 km) y órbita alta (más de 36000 km). Luego, determinaron que error en predicciones de hasta 7 días en el futuro puede ser de aproximadamente 10 kilómetros para los satélites de órbita baja, 1 kilómetro para las órbitas medias y 10 kilómetros para las de órbita alta.

Esta precisión es suficiente para los cálculos necesarios de este trabajo. De acuerdo con la información de la NASA, un objeto a unos 242km de altura debe mantener una velocidad orbital de unos 27355 kilómetros por hora o unos 7.5

kilómetros por segundo. Por lo que, una diferencia de 5 kilómetros en la predicción de la órbita corresponde a menos de un segundo de diferencia en la observación.

2.4 Ingeniería de datos

La definición básica o pura de ingeniería según indica (Yirda, 2019), es una profesión en la que los conocimientos científicos y empíricos se aplican para la conversión óptima de los materiales y fuerzas de la naturaleza en usos prácticos para la humanidad, así como, la invención, perfeccionamiento y utilización de la técnica industrial, y a la resolución de problemas técnicos-sociales. Por tanto, este concepto llevado a la ingeniería de datos nos da como resultado una profesión en donde la aplicación de conocimientos científicos y empíricos se utilizan para la manipulación y conversión de los datos en un producto utilizable para la humanidad, en otras palabras, es convertir los datos en información.

Para poder realizar esta tarea, en la ingeniería de datos se hace uso de diversas técnicas que se seleccionan según sea el objetivo a alcanzar, entre estas, el modelado de datos cobra mucha relevancia para esta investigación, ya que es la parte inicial de casi todo proceso de datos.

2.4.1 Modelado de datos

Los datos pueden encontrarse en varias fuentes, procesadores de texto, audios, videos, mensajes y programas en general. A estos se les llama datos crudos, pues no cuentan con una estructura y funcionalidad.

Por ello, “el modelado de datos es una manera de estructurar y organizar los datos para que se puedan utilizar fácilmente por las bases de datos” (Tecnologías- Información, 2018).

Con el modelado de datos, se dan las instrucciones para crear la base de datos, que brinde la información necesaria, que el cliente requiere. Según Tecnologías- Información, para esta organización de los datos, es importante tomar en cuenta que el modelado de datos, posee tres etapas, a saber:

Modelo de datos conceptuales: son utilizados para explorar conceptos de dominio, con los interesados del proyecto. Se da en una etapa inicial, orienta las reglas y los requerimientos que van a marcar el desarrollo del proyecto.

Modelo de datos lógico: los MDL, permiten la exploración entre los conceptos de dominio y la relación con los dominios del problema.

Modelo de datos físico: se utiliza para diseñar el esquema interno de la base de datos, que representa las tablas de datos, las columnas de datos de esas tablas y las relaciones entre las tablas.

2.4.2 ETL

Los ETL, son procesos de extracción, transformación y carga de datos. De acuerdo con el sitio Power Data (2020), “son una parte de la integración de datos, pero es un elemento importante cuya función completa el resultado de todo el desarrollo de la cohesión de aplicaciones y sistemas”

Mediante este proceso, los datos son extraídos y procesados en otra base de datos que permita su análisis, en función de lo que se desea conocer. Para el sitio, Power Data, existen 2 formas básicas para extraer los datos:

Acumulación simple: consiste en realizar un resumen de todas las transacciones comprendidas en el período de tiempo seleccionado y transportar el resultado como una única transacción hacia el data warehouse, almacenando un valor calculado que consistirá típicamente en un sumatorio o un promedio de la magnitud considerada. Es la forma más sencilla y común de llevar a cabo el proceso de carga.

Rolling: este proceso sería el más recomendable en los casos en que se busque mantener varios niveles de granularidad. Para ello se almacena información resumida a distintos niveles, correspondientes a distintas agrupaciones de la unidad de tiempo o diferentes niveles jerárquicos en alguna o varias de las dimensiones de la magnitud almacenada (por ejemplo, totales diarios, totales semanales, totales mensuales, etc.).

2.4.3 NoSQL

Este término hace referencia a la utilización de bases de datos no estructuradas, que se ha visto desarrollada, con el crecimiento acelerado de las tecnologías, ya que, vienen a flexibilizar el almacenamiento de datos (Graph).

Para la revista Unir (2016), este tipo de base de datos, no es relacional y posee las siguientes ventajas:

- SQL no es el lenguaje de consulta o modificación, sin embargo, si lo soporta.
- Los datos no tienen que almacenarse en tablas.
- Generalmente, almacena en diferentes máquinas, por lo que soporta mejor los posibles fallos, debido a una mayor escalabilidad horizontal.
- Mayor eficiencia en el procesamiento de datos por lo que es más utilizado para aplicaciones que realizan streaming.
- Utilizan la consistencia eventual, lo que significa que los cambios realizados serán replicados en todos los nodos del sistema.

2.4.4 Minería de datos

Según Microsoft Docs (2020) “la minería de datos es el proceso de detectar la información procesable de los conjuntos grandes de datos. Utiliza el análisis matemático para deducir los patrones y tendencias que existen en los datos. Normalmente, estos patrones no se pueden detectar mediante la exploración tradicional de los datos porque las relaciones son demasiado complejas o porque hay demasiado datos”.-

Estos procesos facilitan la gestión de los datos y brindan un aporte fundamental, ya que, de otra forma, no sería posible la utilización masiva de los mismos y los análisis tardarían mucho más tiempo, para llevarse a cabo.

De acuerdo con Microsoft Docs (2020), existen escenarios en los cuales se pueden aplicar la minería de datos:

- **Previsión:** Estimación de ventas, predicción de cargas de servidor o tiempo de inactividad del servidor
- **Riesgo y probabilidad:** Elección de los mejores clientes para el envío de correo directo, determinación del punto de equilibrio probable para los escenarios de riesgo, asignación de probabilidades a diagnósticos u otros resultados

- **Recomendaciones:** Determinación de los productos que es probable que se vendan juntos, generación de recomendaciones
- **Buscar secuencias:** Análisis de las selecciones de los clientes en un carro de la compra y predicción de los siguientes eventos probables
- **Agrupación:** Separación de los clientes o eventos en el clúster de elementos relacionados, análisis y predicción de afinidades

2.4.4.1 Aprendizaje de máquinas

Nick MacCrea (2014) cita a Arthur Samuel en 1959: “Aprendizaje de máquina es el campo de estudio que da a las computadoras la habilidad de aprender sin ser programadas explícitamente.”

Y es que, este tema es uno de los que más movimiento está representando en la actualidad, puesto que es el futuro. A través del mismo, se busca solucionar una serie de problemas que han sobrepasado al ser humano, porque, nuestra capacidad en el manejo de datos masivos, es ciertamente más limitada.

Para ello, Tom Mitchell (citado por MacCrea), indica que “*Se dice que un programa de computadora aprende de experiencia E , con respecto a alguna tarea T y alguna medida de desempeño P , mejora con experiencia E .*”

En la mayoría de las situaciones, que se intentan resolver, la complejidad está presente, y al igual que las personas, vamos resolviendo las cosas de manera más satisfactoria, con la experiencia; las máquinas pueden aprender con la experiencia y con la capacidad de procesar mucha más información.

Para MacCrea, es importante comprender los dos tipos de aprendizaje que se presentan en la resolución de tareas para las máquinas:

Aprendizaje de máquina supervisado: El programa está “entrenado” en un set predefinido de “ejemplos de entrenamiento”, lo cual después facilitó su habilidad de alcanzar una conclusión acertada cuando se le pasa nueva data.

Aprendizaje de Máquina no supervisado: Al programa se le da una cantidad de data y debe encontrar patrones y relaciones dentro de éste.

2.4.4.2 Regresión

La regresión es un modelo matemático, que permite relacionar una variable dependiente, y una o varias variables independientes (Sánchez, 2020). Esto con el fin de comprender, cómo se modifican o influyen las variables estudiadas, para determinar su futuro comportamiento.

Ferre (2019) aporta que, *“la regresión es el conjunto de técnicas usadas para explorar y cuantificar la relación de dependencia entre una variable cuantitativa llamada variable dependiente o respuesta y una o más variables independientes llamadas variables predictoras”*.

2.4.4.3 Agrupamiento

De acuerdo con Ligdieli (2020) un clúster es la colección de objetos de datos que son similares entre sí dentro del mismo grupo, clase o categoría y son diferentes de los objetos de los otros clústeres.

La agrupación es una técnica de Aprendizaje no Supervisada en la que hay clases predefinidas e información previa que define cómo se deben agrupar o etiquetar los datos en clases separadas. También podría considerarse como un proceso de análisis exploratorio de datos que nos ayuda a descubrir patrones ocultos de interés o estructura en los datos.

La agrupación en clúster también puede funcionar como una herramienta independiente para obtener información sobre la distribución de datos o como un paso de preprocesamiento en otros algoritmos.

2.4.4.4 Red Neuronal

Las redes neuronales artificiales son una técnica de minería de datos que simula el funcionamiento de las neuronas en el cerebro.

Una red neuronal es un modelo simplificado que emula el modo en que el cerebro humano procesa la información: Funciona simultaneando un número elevado de unidades de procesamiento interconectadas que parecen versiones abstractas de neuronas.

Las unidades de procesamiento se organizan en capas. Hay tres partes normalmente en una red neuronal: una capa de entrada, con unidades que representan los campos de entrada; una o varias capas ocultas; y una capa de salida, con una unidad o unidades que representa el campo o los campos de

destino. Las unidades se conectan con fuerzas de conexión variables (o ponderaciones). Los datos de entrada se presentan en la primera capa, y los valores se propagan desde cada neurona hasta cada neurona de la capa siguiente. al final, se envía un resultado desde la capa de salida.

La red aprende examinando los registros individuales, generando una predicción para cada registro y realizando ajustes a las ponderaciones cuando realiza una predicción incorrecta. Este proceso se repite muchas veces y la red sigue mejorando sus predicciones hasta haber alcanzado uno o varios criterios de parada.

Basogain (1998), explica que las redes neuronales “están constituidas por elementos que se comportan de forma similar a la neurona biológica en sus funciones más comunes. Estos elementos están organizados de una forma parecida a la que presenta el cerebro humano.”

Este autor manifiesta que las redes neuronales poseen características propias del cerebro humano, a saber:

Aprender: adquirir el conocimiento de una cosa por medio del estudio, ejercicio o experiencia. Las ANN pueden cambiar su comportamiento en función del entorno.

Generalizar: extender o ampliar una cosa. Las ANN generalizan automáticamente debido a su propia estructura y naturaleza.

Abstraer: aislar mentalmente o considerar por separado las cualidades de un objeto.

Capítulo 3. Marco Metodológico

3.1 Tipo de Investigación

El tipo de investigación es aplicada, porque se ajusta a las características del proyecto. Este busca atender la necesidad particular de observatorios profesionales en la región centroamericana.

“La investigación aplicada tiene como propósito primordial la solución de problemas [...] Muchos problemas prácticos necesitan respuestas inmediatas que los investigadores básicos no tienen. La necesidad de una respuesta a un problema práctico es una razón legítima para realizar la investigación y puede ser satisfactoria, especialmente si los resultados tienen un impacto inmediato en el mundo.” (Martin & Dávila, 2008)

El presente trabajo busca hacer una estimación y exponer la situación actual de la región en cuanto a la cantidad de satélites y qué tanto pueden estos afectar la observación astronómica, basados en datos del tránsito, manteniendo una base de datos actualizada que monitoree continuamente las fuentes de información en busca de actualizaciones que produzcan un cambio en el panorama actual.

Para esto, se utilizan datos existentes proveídos en bases de datos públicas y se implementan algoritmos ya desarrollados y ampliamente explorados en el contexto de la ingeniería de datos y rastreo satelital.

3.2 Alcance Investigativo

El alcance investigativo es principalmente descriptivo, ya que se utiliza información tránsitos de satélites y se recolecta información sobre ellos para mostrar la afectación en la zona a estudiar.

Para Serrano (2010) este método: “Describe, analiza, registra e interpreta las condiciones que se dan en una situación y momento determinado. Generalmente se pueden contrastar situaciones o hechos, pretendiendo encontrar relaciones causa-efecto entre variables existentes, aunque no manipuladas”

También se plantea un alcance exploratorio debido a que, aunque ya se han realizado estudios que miden la afectación de los satélites en otras latitudes, ninguna ha utilizado una metodología y enfoque iguales a los planteados en este trabajo.

“El objetivo principal de la investigación exploratoria es captar una perspectiva general del problema. [...] La investigación exploratoria se puede aplicar para generar el criterio y dar prioridad a algunos problemas. El estudio exploratorio también es útil para incrementar el grado de conocimiento del investigador con respecto al problema. Especialmente para un investigador que es nuevo en el campo del problema.” Namakforoosh (2005)

3.3 Enfoque

Para este proyecto se adoptó el enfoque alternativo, ya que el mismo brinda la capacidad de poder explicar los postulados epistemológicos, ontológicos y axiológicos de la investigación científica, mismos que se aplican estrechamente a la tecnología y específicamente a la investigación informática.

Este enfoque además se elige en un contexto de disponer una metodología práctica que agilice la investigación y permita realizar el análisis de los datos y su respectiva interpretación de una manera flexible, donde las metodologías tradicionales carecen de una aplicación clara para tal motivo, como así lo expresa Naranjo (2020): “En ese sentido, las fuentes tradicionales para impartir cursos de metodología de investigación muestra una clara carencia para la profesión informática. Los paradigmas que más claramente se abarcan son el positivista y el paradigma naturalista, con sus respectivos enfoques cuantitativo y cualitativo, además de una fuerte inclinación hacia las particularidades de las ciencias sociales y de las humanidades.”

Naranjo (2020) también menciona que: “utilizar un enfoque adecuado tiene entre sus ventajas que al sumar el pragmatismo desde el enfoque alternativo permiten al investigador moverse con libertad por entre una amplia gama de diseños de investigación, así como los instrumentos que considere adecuados para la consecución de sus objetivos.”

En cuanto a la epistemología, la presente investigación procede a estudiar el impacto que tiene el brillo reflejado por los satélites sobre las observaciones astronómicas, además se recolectarán estas estadísticas para poder dimensionar el impacto a través del tiempo. Los investigadores no son parte, ni alteran el objeto de estudio, sino que, tienen un papel de observadores.

En la dimensión ontológica, Gruber (1993) define la ontología de la siguiente manera: “En el contexto de las ciencias informáticas y de la información, una ontología define un conjunto de primitivas de representación con las que modelar un dominio de conocimiento o discurso. Las primitivas de representación suelen ser clases (o conjuntos), atributos (o propiedades) y relaciones (o relaciones entre los miembros de clase).”

Es decir, debe poderse representar el conocimiento obtenido de una forma entendible y básica.

La representación ontológica del objeto de estudio se define en la Figura 11.

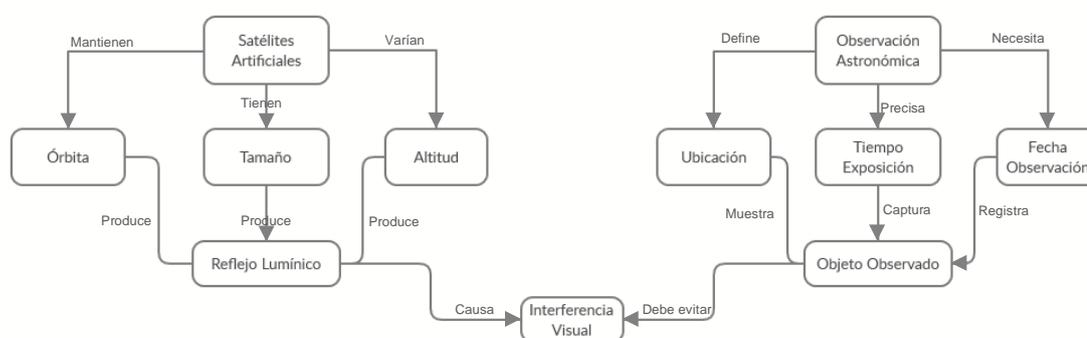


Figura 11: Representación ontológica de la investigación. Fuente: elaboración propia

La investigación pretende demostrar el impacto del reflejo lumínico de los satélites artificiales en las observaciones astronómicas en un punto específico.

La dimensión axiológica de la investigación busca dar una valoración a ese impacto visual basado en la magnitud del brillo de los satélites en una ubicación y hora específica y se categorizará en Bajo, Medio y Alto. Para la realización de esta escala se utiliza como base lo expuesto por Hainaut y Williams (2020).

Clasificación	Factores (satélite individual)			
	Elevación del sol sobre el horizonte	Elevación del satélite en el cielo	Magnitud aparente	Tiempo de exposición (observación)
No relevante	N/A	el. < 0°	?	N/A
Baja	el. < -18°	0° < el. < 15°	mag ≥ 5	ex. ≤ 1s
Media	-18° ≤ el. < -6°	15° < el. ≤ 30°	0 < mag < 5	1s < ex. ≤ 100s

Alta	$-6^\circ \leq \text{el.} < 0^\circ$	$30^\circ < \text{el.}$	$\text{mag} \leq 0$	$\text{ex.} \geq 1000\text{s}$
------	--------------------------------------	-------------------------	---------------------	--------------------------------

Tabla 10: Escala para la clasificación del impacto en la observación astronómica. Fuente: Elaboración propia

Para la determinación de la clasificación en esta escala se utilizan los factores a descritos en la Tabla 10 y se hace un recuento de la cantidad de factores que caen en cada categoría. Si un satélite cumple con condiciones de diferentes clasificaciones, el modelo ponderará el resultado final de todos los factores y hará la determinación basado en una nota entre 0 y 1. Como se muestra en la Tabla 11.

Nivel de afectación	Resultado ponderado del modelo
No Relevante	Resultado entre 0 y 3
Bajo	Resultado entre 4 y 6
Medio	Resultado entre 7 y 9
Alto	Resultado entre 10 y 12

Tabla 11: Determinación de clasificación por nota ponderada. Fuente: Elaboración propia

Es importante mencionar que, para la simplificación de los modelos, a diferencia de Hainaut y Williams (2020) no se toma en cuenta como factor el campo de visión del telescopio, ya que este puede ser aplicable a un caso en específico y no generalizable a más observatorios en la región.

3.4 Diseño

Dada la naturaleza de la investigación, el tipo de diseño que se va a aplicar es de investigación evaluativa.

“En la investigación evaluativa, el investigador suele estar con frecuencia contratado por un agente para describir y evaluar un programa de cambio, con el fin de mejorarlo o de suprimirlo. El resultado de tal investigación normalmente consiste en un informe escrito.” (Cook, Reichardt, & Manuel, 2005)

Si bien es cierto, esta investigación no pretende realizar una evaluación con el fin de cambiar el panorama actual, el informe final tiene como meta alertar a la institución interesada sobre la situación existente, para así ser una base en la toma de decisiones para la planificación en el uso de los recursos y medidas para contrarrestar los efectos negativos.

Para esto se procede a definir los siguientes pasos:

- Obtener los datos actuales e históricos desde el sitio Spacetrack.org

- Almacenar los datos y aplicar procesos de limpieza.
- Se aplicarán algoritmos de ubicación satelital para obtener la trayectoria del satélite a futuro e identificar si causará alguna obstrucción en el cielo.
- Se emplearán algoritmos de clasificación para agrupar los satélites según su nivel de impacto en las observaciones.
- Emplear modelos de aprendizaje automático para asignar valores faltantes a las características de los satélites basados en otros similares.
- Desarrollar modelos de aprendizaje automático con el fin de presentar los escenarios futuros.
- Diseñar un reporte donde se visualicen las métricas mencionadas anteriormente.

3.5 Población y Muestreo

El presente trabajo no implementa conteos estadísticos en el que se apliquen encuestas o cuestionarios a sujetos de estudio, dado que esta investigación no persigue un enfoque cuantitativo.

3.6 Instrumentos de Recolección de Datos

Para conseguir los datos necesarios para la presente investigación se utilizará las interfaces de aplicación (API) del sitio web Space-Track.org para la descarga de los datos TLE los cuales son insumo base para el cálculo de las órbitas. También se contará con información sobre la magnitud standard de los satélites, obtenida del sitio web Satlfare.com.

Se planea utilizar información de todos los satélites registrados que se encuentren orbitando la tierra. Los datos de ubicación a utilizar serán de máximo un año de antigüedad.

Todos los datos mencionados anteriormente son de acceso público, sin embargo, los obtenidos desde Space-Track.org requieren de la creación de un usuario para la utilización de los mismos.

Para una mejor ilustración de los datos se muestra una línea de datos real de un elemento de 2 líneas de la estación espacial internacional:

ISS (ZARYA)

1 25544U 98067A 08264.51782528 -.00002182 00000-0 -11606-4 0 2927

2 25544 51.6416 247.4627 0006703 130.5360 325.0288 15.72125391563537

Y la Figura 12 presenta la información referente a la magnitud estándar (o intrínseca) del satélite, necesaria para el cálculo de la magnitud aparente.

Basic Satellite Search

Search Satellites by Name (or part) Sort by Order

Filter by Launch Year

Quick search of the last 20 objects

Search Results
Click on the satellite name to track it.
Click on the International Designator to view all the other objects launched during that launch.
Click on the Year to view all the other objects launched during that year.

N#	SATELLITE NAME	NORAD	INT. DES.	LAUNCH YEAR	STD MAG	RCS	LAST TLE EPOCH
1	STARLINK-41	44277	19029AU	2019	5.5	3	20221.26324916

Found 1 satellite.

Figura 12: Ejemplo de magnitud estándar tomada de: <http://www.satflare.com/search.asp?sat=STARLINK-41>

3.7 Técnicas de Análisis de Información

Se utilizará el diagrama de flujo de datos para ilustrar los pasos a llevar para la extracción y el análisis de la información proveniente de las fuentes mencionadas anteriormente. En la Figura 13 se detalla el flujo desde la captura de los datos hasta la evaluación final.

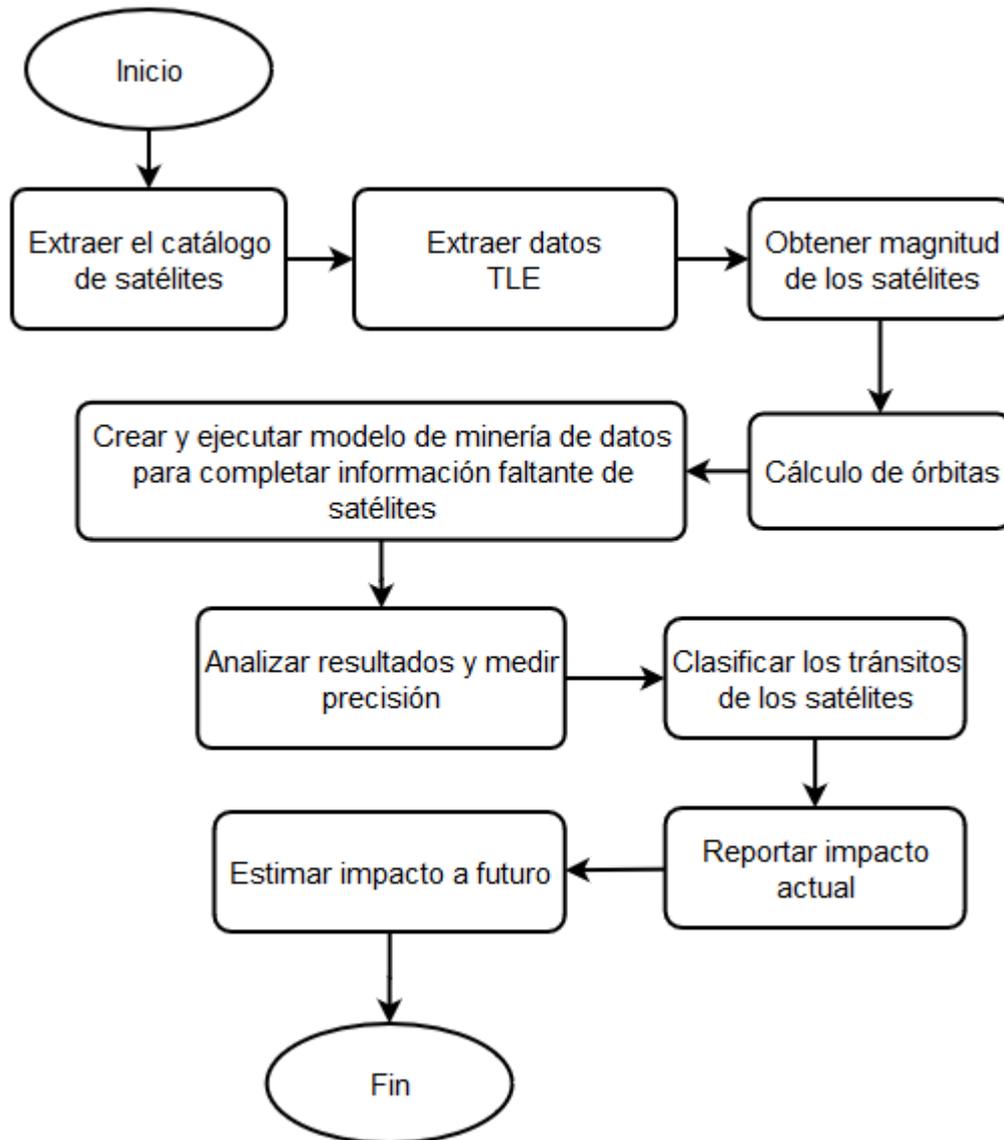


Figura 13: Diagrama de flujo del proceso de obtención y análisis de la información. Fuente: Elaboración propia

Capítulo 4. Análisis del Diagnóstico

A continuación, se realiza un análisis de los datos obtenidos en la sección 3.6. Esto incluye 3 sets de datos los cuales son: el catálogo de satélites, los datos TLE y el catálogo de satélites de Satflare.com, el cual incluye la magnitud intrínseca para los satélites, la cual es utilizada para calcular la magnitud aparente.

Todos estos datos fueron almacenados en formato de texto y descargados manualmente de los sitios fuente.

4.1 Catálogo de satélites

El sistema de space-track rastrea objetos que superan los 10cm de diámetro esto incluye entre otros, basura espacial, escombros de colisiones, satélites y cuerpos de cohetes utilizados para llevar la carga.

La descripción de los datos es la mostrada en la Tabla 12

Nombre del campo	Descripción
INTLDES	Código de satélite internacional
NORAD_CAT_ID	Código de satélite de NORAD
OBJECT_TYPE	Tipo de objeto
SATNAME	Nombre del Satélite
COUNTRY	País de origen del satélite
LAUNCH	Fecha de lanzamiento
SITE	Lugar de lanzamiento
DECAY	Fecha de salida de órbita (final de operación)
PERIOD	Minutos que el objeto tarda en completar una órbita completa
INCLINATION	Inclinación respecto al plano ecuatorial
APOGEE	Punto de la órbita donde el satélite está más alejado de la tierra
PERIGEE	Punto de la órbita donde el satélite está más cercano a la tierra
COMMENT	Información adicional del satélite
COMMENTCODE	Código de comentario
RCSVALUE	Tamaño del corte transversal de radar (fuera de uso)
RCS_SIZE	Categoría basada en el tamaño del corte transversal de radar
FILE	(No especificado)
LAUNCH_YEAR	Año de lanzamiento
LAUNCH_NUM	Número de lanzamiento

LAUNCH_PIECE	Identificador de objeto
CURRENT	Indica si el registro es el más reciente
OBJECT_NAME	Nombre del satélite
OBJECT_ID	Código de satélite internacional
OBJECT_NUMBER	Código de satélite de NORAD

Tabla 12: Diccionario de datos del catálogo de satélites. Fuente: Elaboración propia

Por consiguiente, esta información debe ser filtrada para incluir solamente aquellos objetos que sean denominados como “carga útil”, esto incluye satélites, telescopios espaciales y la ISS que estén orbitando de la Tierra. Esta selección se hace con el fin de mantenerse dentro del alcance de la investigación.

Dadas las restricciones existentes en cuanto al reenvío de información, los investigadores del presente trabajo no pueden compartir el set de datos, sin embargo, a la fecha de realizado el mismo la Tabla 13 muestra el recuento de satélites filtrados y el motivo del filtro.

Descripción	Objetos rastreados
Cantidad total de objetos rastreados por NORAD	46466
Escombros de colisiones	30538
Cuerpos de cohetes	6182
Otros a ser identificados	160
Objetos de carga útil que ya no están orbitando	3529
Objetos en el espacio, pero no orbitan la tierra actualmente	251
Objetos de carga útil (payload)	5806

Tabla 13: Detalle de objetos rastreados por SpaceTrack.org. Fuente: Elaboración propia

En la primera línea de la Tabla 13 se toma el total de objetos que son y han sido rastreados por NORAD (Mando Norteamericano de Defensa Aeroespacial). En la segunda línea aquellos que son considerados como escombros, ya sean resultado de colisiones o basura espacial. Los cuerpos de cohetes están en la tercera línea y estos no son tomados en cuenta para la presente investigación dado que estos se mantienen en órbita por periodos de tiempo cortos ya que hacen su reentrada a la tierra una vez la carga es desplegada.

Del set de datos también se eliminan aquellos objetos que ya fueron decaídos y no orbita más la tierra, ya sea por una reentrada o porque fueron empujados fuera de la órbita terrestre en sus últimos momentos de servicio.

Por último, se filtran los objetos que fueron rastreados en su lanzamiento pero que su destino no era orbitar la tierra, si no misiones a otros puntos del sistema solar, como lo son la Luna y Marte.

4.2 Elementos de dos líneas (TLE)

Este conjunto de datos varía en gran medida, ya que se actualiza la información de cada satélite en momentos diferentes del día. Por lo que para efectos de la investigación, se tomará solamente el último registro disponible para cada satélite en el momento de la extracción.

Para esto, space-track pone a disposición una API la cual retorna el último TLE conocido para el satélite. Para los satélites a utilizar en la investigación el set de datos devuelve un total de 5590 registros, con un tamaño aproximado de 2.5MB.

Esto indica que existen satélites que no se han rastreado o para los cuales la información no está disponible, por tanto, no podrán ser tomados en cuenta para fines del reporte final.

La descripción detallada de cada uno de los atributos del conjunto de datos se muestra en la Tabla 14.

Nombre del campo	Descripción
ORDINAL	Número de registro devuelto
COMMENT	Comentario del origen de datos
ORIGINATOR	Programa utilizado para generar la información
NORAD_CAT_ID	ID de satélite de NORAD
OBJECT_NAME	Nombre del satélite
OBJECT_TYPE	Tipo de satélite
CLASSIFICATION_TYPE	Clasificación del satélite
INTLDES	Identificador internacional
EPOCH	Época del TLE
EPOCH_MICROSECONDS	Microsegundos de la época
MEAN_MOTION	Velocidad orbital media

ECCENTRICITY	Excentricidad
INCLINATION	Inclinación
RA_OF_ASC_NODE	Asención recta del nodo ascendente
ARG_OF_PERICENTER	Argumento del perigeo
MEAN_ANOMALY	Anomalía media
EPHEMERIS_TYPE	Tipo de efeméride
ELEMENT_SET_NO	Número de elemento
REV_AT_EPOCH	Número de revolución
BSTAR	Coefficiente de arrastre utilizado en los algoritmos SGP/SGP4
MEAN_MOTION_DOT	Primera derivada de la velocidad orbital media
MEAN_MOTION_DDOT	Segunda derivada de la velocidad orbital media
FILE	No especificado
TLE_LINE0	Línea 0 del elemento
TLE_LINE1	Línea 1 del elemento
TLE_LINE2	Línea 2 del elemento
OBJECT_ID	Código de satélite internacional
OBJECT_NUMBER	ID de satélite de NORAD
SEMIMAJOR_AXIS	Semieje mayor
PERIOD	Periodo
APOGEE	Apogeo
PERIGEE	Perigeo
DECAYED	Indica si el satélite está fuera de servicio

Tabla 14: Atributos en conjunto de datos de TLE. Fuente: Elaboración propia

4.3 Magnitudes estándar

La magnitud aparente de los satélites será calculada utilizando como insumo, entre otros, la magnitud estándar reportada públicamente por la comunidad de observadores en la página web Satflare.com como se mencionó en la sección 3.6

Para el presente trabajo inicialmente, se supuso la existencia de este valor para al menos el 80% de los satélites en tránsito, sin embargo, los hallazgos desmintieron este supuesto.

Este conjunto de datos reporta un total de 7392 satélites, de los cuales solamente 1677 tienen asignada la propiedad de magnitud aparente, lo cual representa un 22% del total de la base. Es importante destacar que estos 7392 satélites no corresponden todos a satélites activos en órbita, sino que también considera registros de satélites que ya fueron puestos fuera de la órbita terrestre.

Dado que el catálogo de satélites no provee mayores detalles de las características físicas, o materiales del cuerpo del satélite, los investigadores solamente tienen un recurso que caracterice a los cuerpos; el valor RCS (Corte Transversal de Radar), el cual indica de manera aproximada cuál es el tamaño del satélite.

Por lo tanto, para la elaboración de esta investigación se decidió asignar un promedio en el valor de magnitud estándar a aquellos satélites que no lo poseen, categorizando los satélites por el tamaño RCS reportado por SpaceTrack, de aquellos que sí lo tengan disponibles. El procedimiento utilizado para este cálculo se detalla en la sección **5.3.6.1 Cálculo de la magnitud aparente faltante**.

El valor de magnitud aparente es uno los datos utilizados para categorizar el impacto en bajo, medio o alto, siguiendo las pautas propuestas por Hainaut & Williams (2020) tal como se menciona en la sección 3.3.

Capítulo 5. Propuesta de Solución

5.1 Lenguaje de programación y bibliotecas a utilizar

Para el presente trabajo se seleccionó Python como lenguaje de programación principal. Este lenguaje generalmente se puede utilizar en ambientes con una infraestructura básica, permitiendo bajar los costos de implementación y mantenimiento.

Otra razón importante para esta elección es la existencia de la librería Skyfield, una de las mayores librerías disponibles que utiliza los algoritmos propagadores del modelo SGP4/SDP4. Esta librería es utilizada para calcular la posición de las estrellas, planetas y satélites orbitando la tierra. Esto provee una capa de abstracción a los investigadores, quienes pueden remitirse utilizar los cálculos existentes y evitar desarrollar algoritmos propios que incrementen la complejidad del trabajo en cuestión.

5.2 Infraestructura de hardware y tecnologías de almacenamiento

En cuanto a la infraestructura necesaria para el desarrollo del proyecto, se optó por una plataforma como servicio. Por ende, ni los investigadores ni los interesados en dicha solución, deben mantener ni ensamblar una arquitectura adicional para el procesamiento de los datos, si no que se utilizan los recursos de Microsoft, mediante su plataforma en la nube Azure.

Luego de realizar el análisis de los datos y el costo de las diferentes alternativas basadas en la nube, se determinó que la mejor tecnología de base de datos para almacenar la información de las fuentes mencionadas es una base de datos SQL Azure (sin servidor). La cual presenta la ventaja de ser libre de mantenimiento, tiene un costo relativamente bajo y fácilmente integrable con los sistemas de extracción y carga de información.

Para la visualización de los datos, se decidió utilizar Power BI Desktop, el cual puede ser utilizado de manera gratuita y no requiere licenciamiento. Pero deja abierta la posibilidad a cualquier individuo o institución que siga y replique este proceso, pueda publicar los datos en línea y que sean accedidos por su organización.

Azure provee todas las tecnologías necesarias para el procesamiento y almacenamiento de los datos sin tener que invertir más de aproximadamente \$35 USD mensuales para mantener el sistema en ejecución, (valor estimado al momento de realizar la investigación) como se observa en la Figura 14.

Recurso	Costo	Resource type	Location	Resource group name
> tfgsat / tfgsat	\$14.56	SQL Server	us west	personalrg
> tfgbatchaccount / windows4cores	\$11.35	Cuenta de Batch	us west	personalrg
> tfg-adf	\$8.87	Data factory (V2)	us east 2	personalrg
> tfgml9067441144	\$0.02	Cuenta de almacenamiento	us west	personalrg
> tfgbatchaccount / windows4cores	<\$0.01	Cuenta de Batch	intercontinental	personalrg

Figura 14: Desglose de costos mensuales de los recursos de infraestructura. Fuente: Elaboración propia.

La Figura 14 muestra la distribución en los costos, siendo la base de datos Azure SQL el recurso más costoso (\$14.56). Seguido por la cuenta Batch (\$11.35) encargada de crear las máquinas virtuales de manera dinámica, para la ejecución el código de Python del cálculo de los tránsitos, extraer los datos de Satflare y crear las predicciones. También se observa el costo de ejecución del Azure data Factory (\$8.87), el cual se encarga de orquestar la ejecución y mover los datos. Y por último el costo relacionado a la cuenta de almacenamiento, la cual es prácticamente inexistente con solamente 2 centavos de dólar.

5.3 Extracción, transformación y carga de los datos

5.3.1 Conjunto de datos base

Para almacenar los tránsitos y calcular el impacto (alto, medio, bajo) que cada satélite tendría se utiliza la tabla SatEvents la cual contiene los atributos que se consideraron necesarios esta clasificación. En la Tabla 15 se detalla el diccionario de cada uno de los campos.

Columna	Tipo de dato	Descripción
EVENTID	Entero	Id Autogenerado
SATID	Entero	ID NORAD

CITY	Carácter	Ciudad utilizada como referencia para los tránsitos
UTCCALCDATE	Fecha	La fecha del cálculo (UTC)
TLEEEPOCH	Fecha	La fecha del TLE (Campo EPOCH)
UTC0R	Fecha	Fecha y hora (UTC) en la que el satélite se eleva por encima del horizonte
UTC0S	Fecha	Fecha y hora (UTC) en la que el satélite se oculta por debajo del horizonte
UTC15R	Fecha	Fecha y hora (UTC) en la que el satélite llega a 15°
UTC15S	Fecha	Fecha y hora (UTC) en la que el satélite se pone por debajo de 15°
UTC30R	Fecha	Fecha y hora (UTC) en la que el satélite llega a 30°
UTC30S	Fecha	Fecha y hora (UTC) en la que el satélite se pone por debajo de 30°
UTCMAX	Fecha	Fecha y hora (UTC) en la que el satélite llega a su máxima elevación
MAXELEV	Decimal	Máxima elevación (en grados) que el satélite alcanza durante su tránsito
SUNELEVAT0R	Decimal	Elevación del sol (en grados) en UTC0R
SUNELEVAT0S	Decimal	Elevación del sol (en grados) en UTC0S
SUNELEVAT15R	Decimal	Elevación del sol (en grados) en UTC15R
SUNELEVAT15S	Decimal	Elevación del sol (en grados) en UTC15S
SUNELEVAT30R	Decimal	Elevación del sol (en grados) en UTC30R
SUNELEVAT30S	Decimal	Elevación del sol (en grados) en UTC30S
SUNELEVATMAX	Decimal	Elevación del sol (en grados) en UTCMAX
ISSUNLIT0R	Binario	Indica si el satélite es iluminado en UTC0R
ISSUNLIT0S	Binario	Indica si el satélite es iluminado en UTC0S
ISSUNLIT15R	Binario	Indica si el satélite es iluminado en UTC15R
ISSUNLIT15S	Binario	Indica si el satélite es iluminado en UTC15S
ISSUNLIT30R	Binario	Indica si el satélite es iluminado en UTC30R
ISSUNLIT30S	Binario	Indica si el satélite es iluminado en UTC30S
ISSUNLITMAX	Binario	Indica si el satélite es iluminado en UTCMAX
UTCSHADOW	Fecha	Fecha y hora en la que el satélite deja o empieza a ser iluminado por el sol durante el tránsito.
ELEVATSHADOW	Decimal	Elevación del satélite en UTCSHADOW
MAG0R	Decimal	Magnitud aparente en UTC0R
MAG0S	Decimal	Magnitud aparente en UTC0S
MAG15R	Decimal	Magnitud aparente en UTC15R
MAG15S	Decimal	Magnitud aparente en UTC15s
MAG30R	Decimal	Magnitud aparente en UTC30R
MAG30S	Decimal	Magnitud aparente en UTC30S

MAGMAX	Decimal	Magnitud aparente en UTCMAX
MAGPRESHADOW	Decimal	Magnitud aparente en UTCSHADOW
AZ0R	Decimal	Ángulo azimutal en UTC0R
AZ15R	Decimal	Ángulo azimutal en UTC15R
AZ30R	Decimal	Ángulo azimutal en UTC30R
AZMAX	Decimal	Ángulo azimutal en UTCMAX
AZ30S	Decimal	Ángulo azimutal en UTC30S
AZ15S	Decimal	Ángulo azimutal en UTC15S
AZ0S	Decimal	Ángulo azimutal en UTC0S

Tabla 15: Estructura de tabla para almacenar los tránsitos. Fuente: Elaboración propia

Cada fila en esta tabla representa un tránsito satelital. Se considera como un tránsito satelital todo evento donde el satélite se eleve por encima del horizonte del observador o expresado de otra forma, que su altitud sea mayor que 0 grados.

Por tanto, esta tabla tiene los atributos necesarios para poder realizar la clasificación del tránsito, tal cual como se plantea en la sección 3.4.

El sistema se diseñó para almacenar esta información de forma diaria, obteniendo un informe de la situación actual para el día en cuestión. Aunque un historial extenso no es necesario para los resultados de la investigación, se consideró que el sistema no tendría problemas al mantenerlo completo para posibles análisis de regresión en el futuro. Por lo que no se añadió ningún procedimiento para purgar dicha información.

En cuanto a la tabla del catálogo de satélites, esta se mantiene en el mismo formato obtenido desde la fuente, con la única excepción de que se agrega un campo adicional en la base de datos llamado "Load_Date" el cual indica la fecha de extracción de los datos para tener un control sobre el monitoreo de las cargas en caso de ser necesario.

5.3.2 Procedimiento de extracción

El procedimiento de extracción de los datos se basa principalmente en consultas HTTP al sitio web spacetrack.org. En este sitio se realizan 2 consultas.

La primera consulta utilizada es para traer información del catálogo de satélites. El cuerpo de la consulta se presenta en la Figura 16. Y en cuanto al

encabezado, este debe tener un atributo para identificar el contenido el cual debe indicar que es tipo "json", como se muestra en la Figura 15.

```
Content-Type: application/json
```

Figura 15: Encabezado de la solicitud HTTP para la extracción del catálogo de satélites. Fuente: Elaboración propia

```
{
  "identity": "usuario de spacetrack",
  "password": "contraseña",
  "query": "https://www.space-
  track.org/basicspacedata/query/class/satcat/OBJECT_TYPE/PAYLOAD/
  orderby/NORAD_CAT_ID asc/format/csv/emptyresult/show
}
```

Figura 16: Código para extracción del catálogo de satélites. Fuente: Elaboración propia

En el cuerpo de la solicitud se detallan 2 atributos que corresponden a la verificación de la identidad (identity y password) el atributo *query* el cual contiene varias partes, el cual se descompone en la Tabla 16.

Parte del URL	Utilidad
https://www.space-track.org/basicspacedata/query/class/	Base del enlace a la API
satcat/	El conjunto de datos a consultar
OBJECT_TYPE/PAYLOAD/	Filtro para excluir todos los objetos que no sean carga útil
orderby/NORAD_CAT_ID asc/	Ordenar el conjunto de respuesta por ID
format/csv/	Retornar la respuesta en formato csv
emptyresult/show	Propiedad opcional, para retornar un conjunto vacío en caso de que no se encuentren objetos

Tabla 16: Desglose del URL para retornar el catálogo de satélites. Fuente: Elaboración propia

La segunda consulta que se utilizó es aquella para extraer información del último elemento de 2 líneas de cada satélite. Sin embargo, es necesario aclarar que no todos los satélites reportan continuamente el estado actual y en otros casos ni siquiera se encuentra información en la base de datos. El cuerpo de la consulta se observa en la Figura 17

```

{
  "identity": "usuario de spacetrack ",
  "password": "contraseña",
  "query": "https://www.space-
track.org/basicspacedata/query/class/tle_latest/ORDINAL/1/NORA
D_CAT_ID/ListIDsCSV/format/csv"
}

```

Figura 17: Código para extracción de TLEs

Esta consulta comparte muchas similitudes con el catálogo de satélites, en cuanto al tipo de llamado y atributos base. En la Tabla 17 se detalla cada parte del atributo *query* el cual representa la llamada a la base de datos.

Parte de la URL	Utilidad
https://www.space-track.org/basicspacedata/query/class/	Base del enlace a la API
tle_latest/	Conjunto de datos a consultar
ORDINAL/1/	Filtra al elemento más reciente
NORAD_CAT_ID/ ListIDsCSV/	Filtra utilizando una lista de ID de satélites
format/csv	Retorna la respuesta en formato csv

Tabla 17: Desglose del URL para retornar los elementos de dos líneas. Fuente: Elaboración propia

En la Tabla 17 se muestra un parámetro llamado ListIDsCSV detallado en color rojo. Esta lista consiste en una serie de identificadores de satélites (NORAD_CAT_ID), separada por coma y cargada en tiempo de ejecución para consultar solamente aquellos elementos que sean válidos en el momento de ejecutar la consulta.

Este proceso se hace en iteraciones de 300 satélites cada 20 segundos. Así, el sistema evita traer más información que la considerada relevante en cada carga y baja la demanda en el servicio de Space-Track.org, el cual tiene restricciones en la cantidad y complejidad de solicitudes por minuto.

El sistema desarrollado utiliza el último elemento de datos existente, sin importar cuán viejo este sea. Los satélites con elementos de datos considerados viejos son la minoría y se detalla cuál es la situación en la Tabla 18.

Categoría	Cantidad de satélites
TLE Reciente	5446
Mas de 24h	218
Sin datos	142

Tabla 18: Distribución de satélites de acuerdo con la antigüedad de TLE

Esto indica que existen 142 satélites sin datos y, por consiguiente, se ignoran. Sin embargo, se tomaron en cuenta los 218 cuales datos existentes tienen una antigüedad mayor a 24 horas, así como aquella mayoría con datos recientes de TLE.

Como ya se comentó en el marco conceptual, las predicciones sobre elementos de dos líneas pierden exactitud con forme más viejos estos sean. Aun así, las mediciones hechas por este sistema no se realizan con la intención de detectar colisiones, si no de poder prever una posible afectación de manera aproximada. Por lo que los autores decidieron mantener la complejidad al mínimo incluyendo estos elementos mientras que se aumenta la cantidad de satélites analizados para una mayor completitud, incluso si la ubicación reportada varía en distancia y ubicación.

Por último, para la extracción de datos de la magnitud aparente. Esta se basa en descargar los datos provenientes de la página Satflare.com mediante una llamada web utilizando el enlace:

<http://www.satflare.com/advsearch.aspx?sat=&ExcDeb=1&ExcRB=1&max=30000>

Este retorna la lista de todos los satélites existentes en su base de datos, con el formato indicado en la Tabla 19.

Columna	Descripción
N#	Número de elemento
NAME	Nombre del satélite
NORAD	Id de NORAD

INT.DESC.	Identificador internacional
PERIGEE[km]	Perigeo en km
APOGEE[km]	Apogeo en km
PERIOD[min]	Periodo mínimo
RCS	Sección radar equivalente
STD.MAG	Magnitud estándar
MAG.	Magnitud aparente
EL.[deg]	Elevación
DEC*[deg]	No especificado
RA*[hour]	No especificado
INCL.[deg]	Inclinación
ECC.	No especificado
AGE[day]	Antigüedad desde última actualización

Tabla 19: Descripción del conjunto de datos de Satflare.com. Fuente: Elaboración propia.

Los datos en este conjunto, según la experiencia obtenida por los investigadores del presente trabajo, no presentan una alta variación. Por ende, para evitar consultar el sitio diariamente, los datos se refrescan cada 2 semanas buscando actualizaciones en los valores de magnitud estándar. Como contingencia, y debido a que no todos los satélites tienen esta información presente se desarrolló un método alternativo para asignar este valor a aquellos satélites donde esta información no esté disponible. Este proceso se muestra en detalle en la sección **5.3.6.**

Estos datos no se almacenan en una tabla con todos sus atributos, ya que algunos de ellos, no son necesarios para los motivos de la investigación. Por lo tanto, la estructura de los datos, está representada en la Tabla 20.

Columna	Descripción
NORAD_CAT_ID	Número de elemento
SATNAME	Nombre del satélite
INTLDES	Identificador internacional
RCS	Tamaño de sección radar equivalente
STDMAG	Magnitud estándar
MAG	Magnitud aparente

LOAD_DATE	Fecha de carga de la información
-----------	----------------------------------

Tabla 20: Descripción de la forma final de los datos de magnitud satelital. Fuente: Elaboración propia

5.3.4 Tecnología de orquestación y cálculo

Como se mencionó en la sección 5.2, este trabajo fue realizado utilizando una tecnología 100% basado en la nube mediante la plataforma Azure de Microsoft.

Esta plataforma provee como servicio Azure Data Factory, la cual es una herramienta para la extracción y carga de datos, así como la orquestación de la ejecución de las tareas.

Por lo tanto, se desarrollaron cinco procesos para la extracción y carga de los datos. Estos procesos son para la carga de: catálogo de satélites, elemento de dos líneas, el cálculo y carga de los tránsitos, magnitud estándar, predicción a futuro y un proceso para ejecutar los demás (orquestador).

Carga del catálogo de satélites

Este proceso de carga utiliza solamente dos actividades. Una para copiar los datos desde un origen HTTP hacia el SQL Server en una tabla temporal. Luego ejecuta un procedimiento almacenado en la base de datos para combinar los datos nuevos con los que fueron extraídos recientemente. Esto como se observa en la Figura 18.

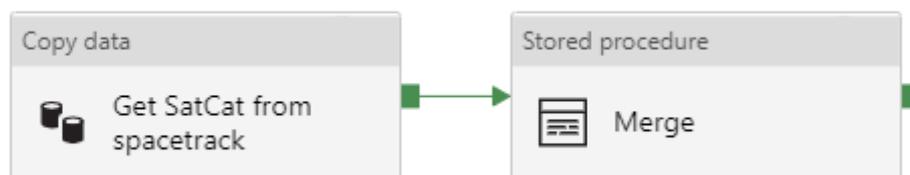


Figura 18: Estructura del proceso de carga del catálogo de satélites. Fuente: Elaboración propia

Carga de los elementos de dos líneas

Este proceso utiliza una estructura particular, dado el tipo de solicitud que hay que realizar al sistema base. Debido a que no se quiso realizar una llamada muy pesada a la base de datos de Space-Track.org, el proceso divide la carga en lotes y extrae un lote cada 20 segundos.

La cantidad de satélites a consultar por lote, se definió en 300, por lo que el número de lotes está dado por la cantidad total de satélites válidos dividido entre el tamaño del lote.

Se utiliza un paquete que consiste en 6 actividades, dos de las cuales se encuentran contenidas en la actividad "Foreach", este siendo un método para iterar por cada ciudad.

En la Figura 19 se observa el proceso, en orden de ejecución, de izquierda a derecha.

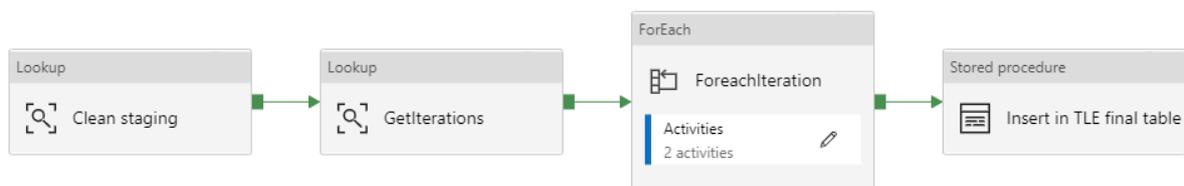


Figura 19: Vista general del proceso de carga de los elementos de dos líneas. Fuente: Elaboración propia.

El mismo, primero limpia la tabla de almacenamiento temporal. Luego, obtiene una lista de lotes calculada desde la base de datos, cada ítem de esta lista corresponde a una cadena de caracteres compuesta por identificadores de satélites separados por coma, los cuales serán consultados.

La tercera actividad se encarga de la iteración y ejecución de cada uno de los lotes. Una vez ejecutado, guarda este fragmento de información en la tabla temporal diseñada para ello y espera 20 segundos antes de la siguiente llamada. La estructura de esta iteración puede verse en la Figura 20.

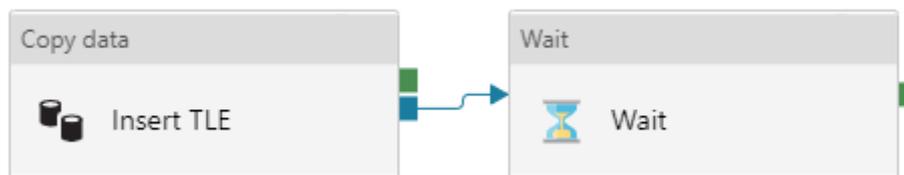


Figura 20: Vista del proceso de ejecución de cada lote de los elementos de dos líneas. Fuente: Elaboración propia

Por último, una vez que todos los lotes han sido ejecutados, se procede con la inserción de la información en la tabla final desde la tabla temporal.

Carga de la magnitud estándar

La carga de la magnitud estándar se realiza en un proceso que incluye tres pasos, los cuales están representados en la Figura 21. El primer paso consiste en la extracción de los datos; la segunda es la carga a una tabla temporal en la base de datos SQL y; la tercera, persistir estos datos en su forma final, detallada en la sección 5.3.

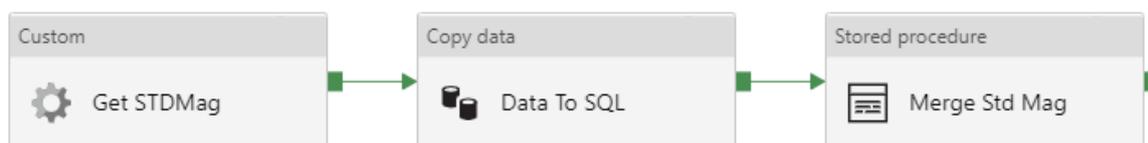


Figura 21: Estructura del proceso de carga de la magnitud estándar. Fuente: Elaboración propia

A diferencia de las cargas anteriores, este proceso necesita realizar la extracción mediante un código programado específicamente para la tarea, ya que está basado en la lectura de código HTML, el cual debe ser pre procesado para su uso.

Por lo dicho, esta tarea se ejecuta desde una máquina virtual, que, gracias a la flexibilidad de Azure, se crea automáticamente para ejecutar el código. Este programa está basado en python y genera como resultado un archivo delimitado por comas, el cual posteriormente es subido a la cuenta de almacenamiento de Azure.

Este archivo es entonces utilizado por la siguiente actividad en la cadena: “Data To SQL”, la cual lee el archivo y se encarga de subirlo a la base de datos en la tabla temporal ya mencionada.

Por último, se ejecuta un procedimiento almacenado que se encarga de borrar los datos antiguos y cargar los nuevos.

5.3.5 Cálculo de los tránsitos

Los tránsitos se calculan de manera diaria luego de extraer la información de los elementos de dos líneas. Cada uno de estos tránsitos es llamado “evento” en el sistema.

El evento, está entonces, definido como una fila de datos en la tabla SatEvents. La cual contiene la información necesaria para calcular el impacto del satélite en un momento dado, durante su tránsito sobre la ciudad estudiada.

El cálculo se realiza con los últimos TLE encontrados para cada satélite. Por tanto, se generan todos los eventos desde las 00:00 hasta las 23:59 horas del día del cálculo.

5.3.5.1 Metodología del cálculo

Como se mencionó anteriormente, estos tránsitos se calculan basados en los elementos de dos líneas. Estos elementos de dos líneas se procesan utilizando el modelo de propagación SGP4/SDP4.

Este modelo al ser bastante conocido y estudiado a través de los años, es de fácil acceso para la comunidad científica. Siendo un proyecto de código abierto, ha sido migrado a diferentes lenguajes. Para este proyecto se utilizó una librería en python llamada Skyfield desarrollada por Brandon Rhodes la cual incorpora varios métodos para simplificar las tareas del cálculo.

La lógica propuesta para la solución se realizó en python utilizando esta librería como recurso principal. La Figura 22 muestra el proceso del cálculo de los tránsitos. El término “magnitud” utilizado representa el brillo del satélite.

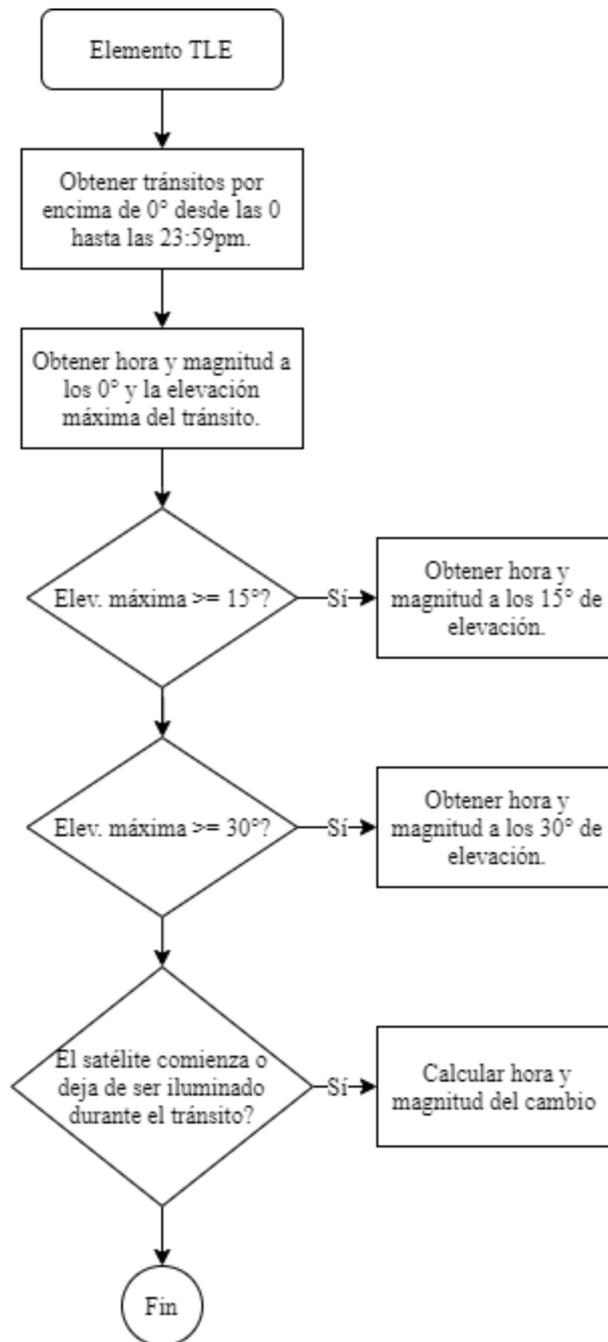


Figura 22: Proceso del cálculo de los tránsitos satelitales. Fuente: Elaboración propia

Este proceso es bastante intensivo en cuanto al uso del procesador. Por lo tanto, se implementó una metodología de multiprocesamiento, donde se aprovecha al máximo cada uno de los núcleos e hilos disponibles en la máquina donde se ejecuta.

La ejecución de este proceso se realiza en una máquina virtual creada de manera dinámica (bajo demanda) en Azure con un procesador de 4 núcleos. Este

cálculo para la cantidad de TLEs actuales, toma alrededor de 10 minutos por país. Y por ende, unos 50 minutos para las cinco capitales centroamericanas.

En la Figura 23 se detalla el proceso en el pipeline de Azure Data Factory.

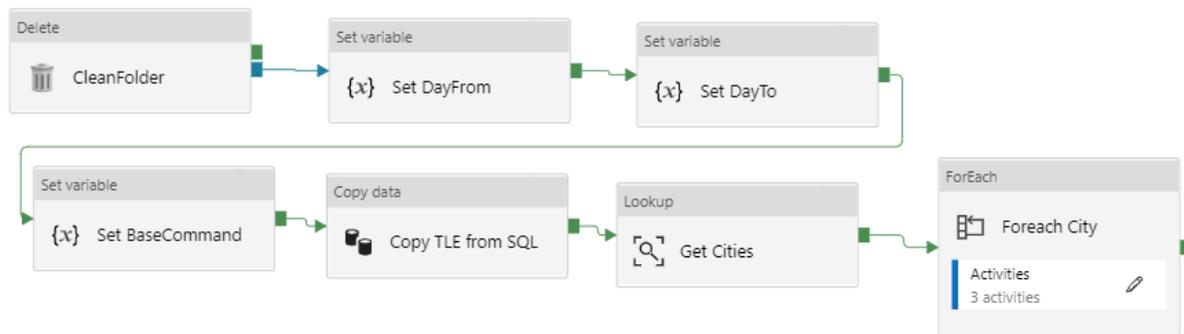


Figura 23: Vista general del proceso de cálculo de los tránsitos. Fuente: Elaboración propia.

Primeramente, se limpian el directorio en la cuenta de almacenamiento de Azure, se determina la fecha inicial y final para los cálculos, luego se extraen los TLEs más recientes hacia el directorio previamente preparado y se extraen los datos de las ciudades a estudiar.

Por cada ciudad se calculan los eventos en la máquina virtual solicitada. Este sube un documento en formato csv a la nube y es posteriormente movido a la base de datos SQL. Como se muestra en las actividades mostradas en la Figura 24

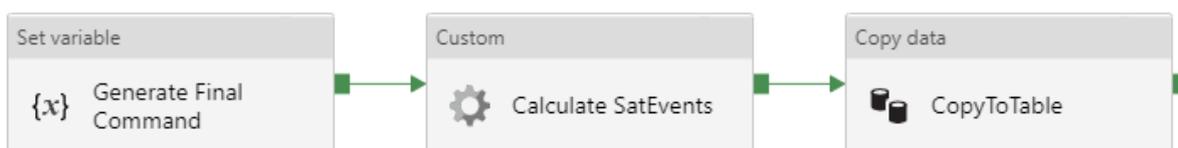


Figura 24: Procesamiento detallado del cálculo de los tránsitos por país. Fuente: Elaboración propia.

5.3.6 Cálculo de la magnitud aparente

Para poder realizar el cálculo estimado de la magnitud aparente de cada satélite, se recurrió a utilizar la fórmula ya mencionada en el marco conceptual. Esta fórmula asume que el objeto, en este caso el satélite, es una esfera difusa que refleja la luz de acuerdo a diferentes características que esta pueda tener en un

momento dado, como lo son el ángulo de incidencia de la luz proveniente del sol, su tamaño, la distancia entre el mismo y la posición del observador.

La fórmula utilizada para determinar la magnitud es la indicada en el marco conceptual. Sin embargo, existe una diferencia en cuanto lo mencionado en la teoría y lo utilizado en la práctica de esta investigación; y es que para este trabajo ya se cuenta con la magnitud intrínseca del satélite utilizando el conjunto de datos de Satflare.

La utilización de este dato simplifica la fórmula ya que, de otra forma, no hubiera sido posible realizar este cálculo sin hacer muchas suposiciones, principalmente porque el tamaño exacto y el tipo de material de cada satélite son desconocidos y no se obtuvo acceso a ninguna base de datos pública que tuviera esta información disponible.

La magnitud intrínseca ya combina el volumen, el albedo y el área de sección transversal, por lo que esto se puede extraer de la fórmula. De esta manera, la fórmula a utilizar está definida como:

$$M_v = M_i - 2.5 \log_{10}(\sin \theta + (\pi - \theta) \cos \theta) + 5 \log_{10}(R)$$

En donde M_i es el valor de magnitud intrínseca. R es la distancia en kilómetros hasta el objeto y θ es el ángulo de fase (ángulo de incidencia formado entre el sol, satélite y observador).

5.3.6.1 Cálculo de la magnitud aparente faltante

Existe una excepción en este cálculo y son aquellos satélites donde la magnitud intrínseca no está disponible. Es por ello que los investigadores recurrieron a utilizar un método de contingencia para poder asignar un valor a los satélites donde se desconoce.

La idea originalmente planteada por el proyecto era la de utilizar un sistema de aprendizaje de máquinas que pudiera responder, basado en características de otros satélites similares, con un valor aproximado de magnitud.

De acuerdo con la fórmula para calcular la magnitud aparente, los valores necesarios son: el albedo, el ángulo de fase, el área de sección transversal, la distancia, la magnitud del sol y el coeficiente de reflejo especular-difuso. Los valores

característicos de cada satélite serían el albedo, el valor exacto del área de sección transversal y el coeficiente de reflejo especular-difuso

Sin embargo, no fue posible acceder a ninguna base de datos que pusiera a disposición estos valores, así como tampoco es viable y sale del alcance de la investigación buscar información de cada satélite de manera manual, sin garantía de encontrar todos los atributos necesarios para el modelo. Sin mencionar que la lista de satélites irá creciendo cada vez más en el tiempo, debiendo así existir la labor de mantener los datos actualizados conforme haya más lanzamientos.

Por lo tanto, dadas las condiciones actuales, los investigadores solamente pudieron utilizar como parámetro el valor de tamaño RCS indicado por space-track.org en su catálogo de satélites para inferir el valor. Así, combinando los datos de Satflare.com y el catálogo ya mencionado, se aplicó un promedio por categoría de RCS (pequeño, mediano y grande). Este promedio se ejecuta antes de realizar el cálculo de los tránsitos, utilizando la información más reciente en la base de datos de todos los satélites que sí posean el atributo de magnitud.

El resultado de cada valor es el mostrado en la Tabla 21. La tabla también muestra la cantidad de satélites que van a utilizar este valor. Para los satélites donde Space-Track.org no indique su tamaño, se tomarán en cuenta como 'Grandes' con el fin de tener en cuenta, de manera conservativa, el peor escenario. Este subconjunto representa un 6% de todos los satélites existentes orbitando la tierra.

Cantidad de Satélites	Categoría	Tamaño Reportado
2922	Sin Magnitud	Grande
881	Con Magnitud	Grande
204	Con Magnitud	Mediano
1106	Sin Magnitud	Mediano
69	Con Magnitud	Pequeño
624	Sin Magnitud	Pequeño

Tabla 21: Conteo de satélites con magnitud intrínseca. Fuente elaboración propia

Por ende, el diagrama de flujo de procesos introducido en la sección 3.7 fue modificado para reflejar la realidad de la implementación. El nuevo diagrama y el proceso afectado puede observarse en la Figura 25.

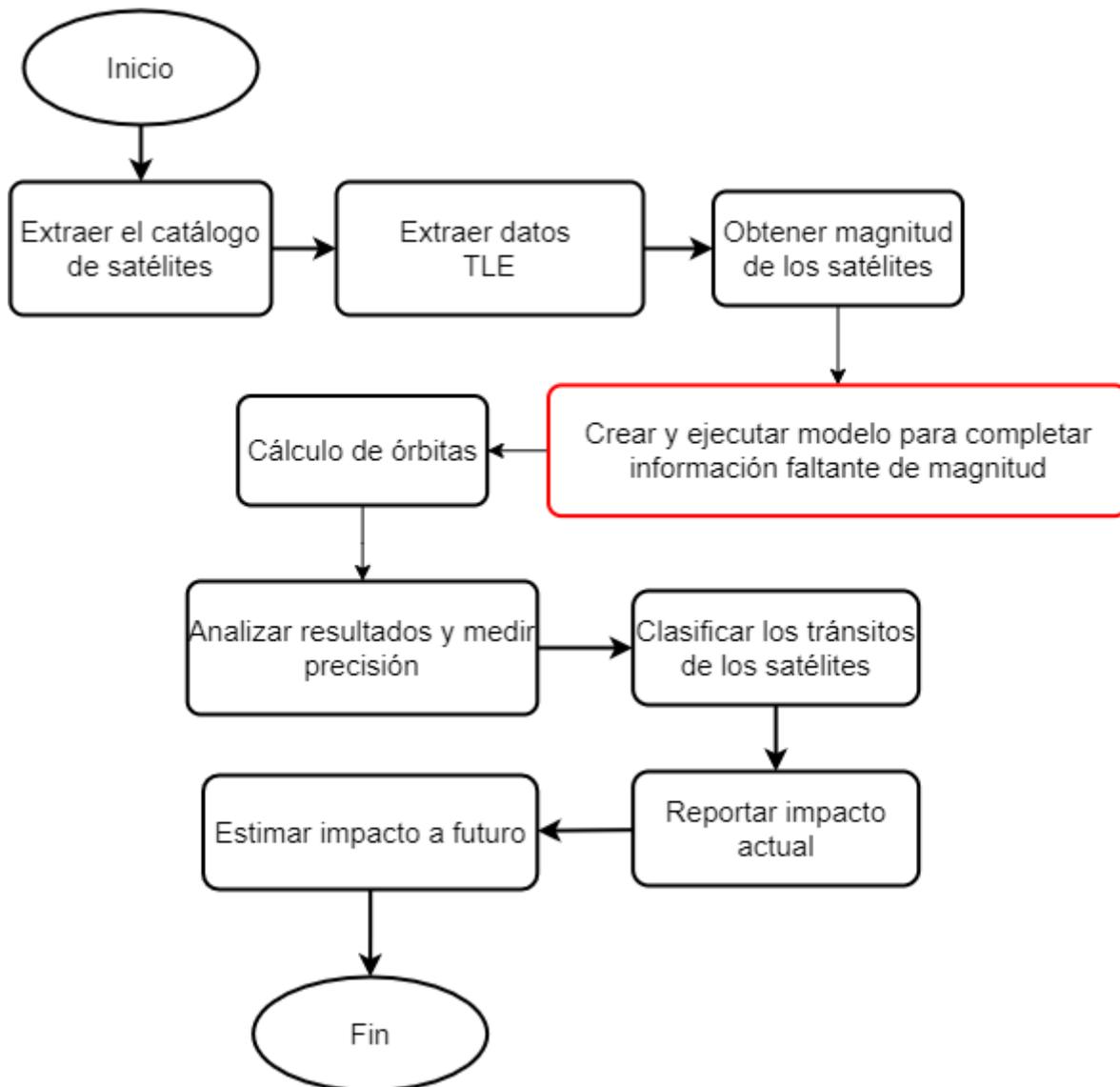


Figura 25: Flujo de procesos modificado. Fuente: Elaboración propia

5.3.7 Clasificación de los tránsitos

Para establecer la clasificación adecuada se realiza la revisión bibliográfica sobre el problema y se encuentra el estudio reciente de Hainaut y Williams (2020), en el cual se establecen los factores de mayor relevancia para el impacto visual de los satélites en las observaciones astronómicas y su respectiva clasificación.

Es por esto que, se considera indispensable tomar en cuenta dicha investigación para el presente proyecto, ya que, orienta el desarrollo del procedimiento de clasificación. Sin embargo, se hace necesario determinar la utilización de esta tabla con los datos obtenidos.

Para cada tránsito satelital se registran diversos factores que permiten obtener un cálculo aproximado de su clasificación de impacto visual, entre estos, los utilizados en el presente estudio para asignar una clasificación son los siguientes

- Magnitud aparente: Se refiere al brillo que emite un satélite.
- Elevación del satélite: Corresponde a la posición en grados del satélite con respecto al horizonte y la posición del observador
- Elevación del Sol: Ubicación del sol en grados con respecto al horizonte y la posición del observador.
- Tiempo de Exposición: Duración en segundos de la captura fotográfica utilizada para la observación astronómica.

Se realiza una valoración de cada factor para cada satélite en el momento de tránsito en sus diferentes ubicaciones y tiempos basándose en la tabla basada en lo expuesto por Hainaut y Williams (2020).

Clasificación	Factores (satélite individual)			
	Elevación del sol sobre el horizonte	Elevación del satélite en el cielo	Magnitud aparente	Tiempo de exposición (observación)
No relevante	N/A	el. $< 0^\circ$?	N/A
Baja	el. $< -18^\circ$	$0^\circ < \text{el.} < 15^\circ$	mag ≥ 5	ex. $\leq 1\text{s}$
Media	$-18^\circ \leq \text{el.} < -6^\circ$	$15^\circ < \text{el.} \leq 30^\circ$	$0 < \text{mag} < 5$	$1\text{s} < \text{ex.} \leq 100\text{s}$
Alta	$-6^\circ \leq \text{el.} < 0^\circ$	$30^\circ < \text{el.}$	mag ≤ 0	ex. $\geq 1000\text{s}$

Figura 26: Escala para la clasificación del impacto en la observación astronómica. Fuente: Elaboración propia

Se le asigna un valor numérico de 0 a 3 a cada factor, según la clasificación correspondiente en donde 0 es Nula, 1 es Baja, 2 es Media y 3 es Alta. Este procedimiento se realiza para poder obtener una clasificación general del impacto que combina todos los factores que participan en un momento dado.

Dado lo anterior, para obtener el valor promedio de todos los factores en un mismo tránsito se suman las clasificaciones y se agrupan los resultados respectivamente a cada categoría de la siguiente forma:

- Valores entre 0 y 3 se clasifican como 0 o impacto Nulo.
- Valores entre 4 y 6 se clasifican como 1 o impacto Bajo
- Valores entre 7 y 9 se clasifican como 2 o impacto Medio
- Valores entre 10 y 12 se clasifican como 3 o impacto Alto

Para efectos de simplicidad en los datos, la hora del tránsito se redondea hacia abajo en segmentos de 5 minutos, es decir, se mostrarán los tránsitos cada 5 minutos iniciando desde el minuto 0 hasta el minuto 55, por ejemplo, si un tránsito ocurre desde la hora 20:17 hasta la hora 20:54 se ubicaran estos datos desde la hora 20:15 cada 5 minutos hasta la hora 20:50. Si los factores llegan a variar en un mismo bloque por ejemplo que un satélite haya pasado de elevación 15° a 30° de la hora 20:32 a la hora 20:34 en este caso se registran ambos eventos en la hora 20:30 y se tomará la clasificación mayor, es decir, la que pudo generar mayor impacto visual. Tal como se muestra en la Figura 27.

City	SATID	Eventid	Fecha	Hora	Minuto	isSunLit	SatElv	SunElv	Magnitud	Sat	Sun	Mag	Obs1s	Obs100s	Obs1000s	Clasificacion
1	SanJose	44348	2425648	2020-11-02	18	0	1	15	-12.9608382335367	9.85514016839454	2	2	1	1	2	2
2	SanJose	44348	2425648	2020-11-02	18	5	1	100	-13.4657636259311	9.09212600896847	3	2	1	2	2	2
3	CiudadGuatemala	46810	2481252	2020-11-02	18	5	1	100	-8.57565734319501	6.8909971410046	3	2	1	2	2	2
4	CiudadGuatemala	46808	2481261	2020-11-02	18	20	1	100	-12.2083986091772	6.74012360587457	3	2	1	2	2	2
5	CiudadGuatemala	46809	2481270	2020-11-02	18	5	1	100	-8.71998167825235	6.88668336014602	3	2	1	2	2	2
6	CiudadGuatemala	46807	2481281	2020-11-02	18	30	1	100	-15.2248353504574	6.5972130207898	3	2	1	2	2	2
7	CiudadGuatemala	46587	2481575	2020-11-02	18	5	1	15	-8.7881588025931	8.10518509126254	2	2	1	1	2	2
8	CiudadGuatemala	46587	2481575	2020-11-02	18	5	1	30	-9.00503840446419	6.57112581462526	2	2	1	1	2	2
9	CiudadGuatemala	46587	2481575	2020-11-02	18	5	1	100	-9.25968466811051	3.52892737980128	3	2	2	2	3	2
10	CiudadGuatemala	46587	2481575	2020-11-02	18	10	1	15	-9.73350632028697	4.97887658320949	2	2	2	2	2	2
11	CiudadGuatemala	46587	2481575	2020-11-02	18	10	1	30	-9.51561158400198	3.98186023671221	2	2	2	2	2	2
12	CiudadGuatemala	46586	2481583	2020-11-02	18	0	1	15	-7.83903258529434	7.13201737385279	2	2	1	1	2	2
13	CiudadGuatemala	46586	2481583	2020-11-02	18	5	1	30	-8.60850535414946	4.13662317235875	2	2	2	2	2	2
14	CiudadGuatemala	46586	2481583	2020-11-02	18	5	1	100	-8.37093654104932	3.97158078425654	3	2	2	2	3	2

Figura 27: Ejemplo de los datos previo a la selección del tránsito de mayor impacto. Fuente: Elaboración propia

Es importante aclarar que, cuando el proceso clasifique alguno de los factores como nulo, se establece que el impacto general combinado con los demás factores también será nulo. Las razones para que un tránsito se clasifique como nulo o “No relevante” son: que suceda durante el día, o que no refleje la luz del sol y por lo tanto no se pueda calcular la magnitud del brillo.

5.3.8 Modelo de predicción de datos

Uno de los principales entregables del presente trabajo es la creación de un modelo de predicción que brinde un vistazo de cómo se ve el futuro si la tendencia de lanzamiento de satélites se mantiene como ha venido dándose a través de los años.

Para lograr esta predicción se utilizó un modelo de series de tiempo basado en dos atributos, el mes/año y la cantidad de satélites en órbita durante ese periodo. Por tanto, la predicción se realiza a nivel mensual.

El lenguaje de programación utilizado para tal fin fue Python con la librería FBProphet, desarrollada y liberada por Facebook para realizar predicciones. La cual reduce la complejidad y aumenta la eficiencia en el desarrollo.

Si se grafica esta información, el resultado a septiembre de 2020 es la mostrada en la Figura 28.

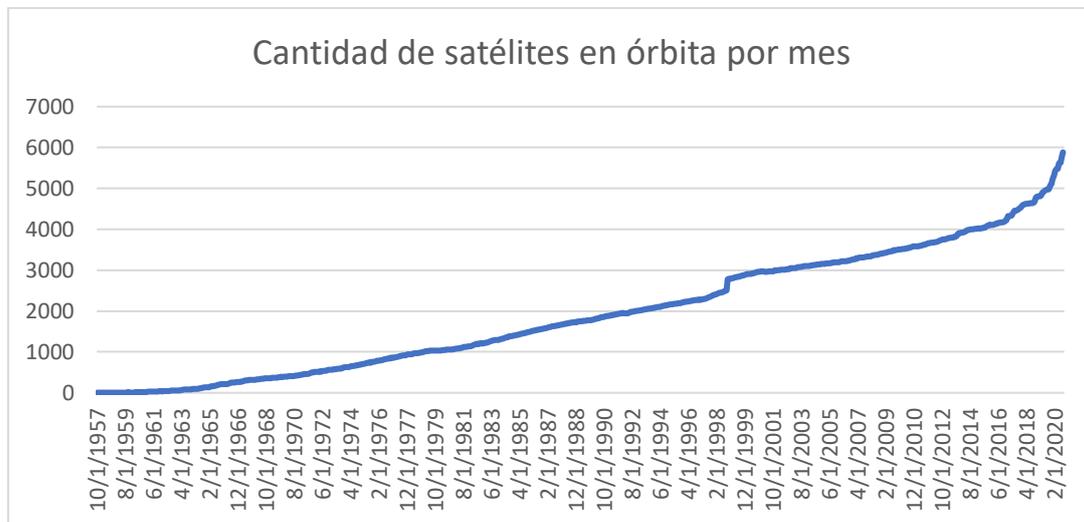


Figura 28: Cantidad de satélites por año. Fuente: Elaboración propia

El gráfico muestra que a finales de 2016 comienza un incremento en la tendencia en cuanto a la cantidad de satélites orbitando la tierra, la cual venía creciendo a una tasa constante.

Este proceso inicia con un set de datos de únicamente 2 columnas, las cuales son el mes y la cantidad de satélites en órbita en el mes en particular.

La librería FBProphet ha sido diseñada para resolver predicciones basadas en series de tiempo, por lo cual la utilización es relativamente sencilla. Solamente es necesario enviar como parámetro un conjunto de datos con las dos columnas mencionadas anteriormente y llamar la función para entrenar el modelo.

Por lo tanto, el set de datos base a utilizar luce como se muestra en la Tabla 22.

Mes Año	Num. Sats.
01/09/2020	5883
01/08/2020	5750
01/07/2020	5622
01/06/2020	5613
01/05/2020	5488

01/04/2020	5488
01/03/2020	5427
01/02/2020	5330
01/01/2020	5229
01/12/2019	5105

Tabla 22: Representación de la preparación del conjunto de datos. Fuente: Elaboración propia.

Una vez que el modelo fue creado y entrenado con los datos ya mencionados anteriormente, se procede a predecir el valor para los siguientes 12 meses tomando como entrada la predicción del mes anterior, tal como se observa en la Tabla 23.

El resultado es un set de datos como el mostrado en la Tabla 23. En el cual se agrega una columna al conjunto de datos original que contenga la predicción realizada para los siguientes 12 meses en el futuro.

Mes Año	Num.Sats	Predicción
01/07/2020	5622	
01/08/2020	5750	
01/09/2020	5883	
01/10/2020		5931
01/11/2020		5980
01/12/2020		6030
01/01/2021		6080
01/02/2021		6132
01/03/2021		6185
01/04/2021		6239
01/05/2021		6293
01/06/2021		6349

01/07/2021		6406
01/08/2021		6465
01/09/2021		6524

Tabla 23: Representación de la predicción de datos. Fuente: Elaboración propia.

Este enfoque solamente tiene un problema, y es que si se utiliza el conjunto de datos en su totalidad la predicción no es realista dado el comportamiento descrito anteriormente sobre el incremento desmesurado de lanzamientos en los últimos 5 años.

La Figura 29 muestra el resultado de la predicción si se utilizara toda la información disponible de lanzamientos desde 1957. Como puede observarse, con esta información la predicción tomaría más de 60 meses para alcanzar los valores actuales. Esto es dado ya que el incremento reciente se sale de la tendencia esperada por el modelo de regresión utilizado.

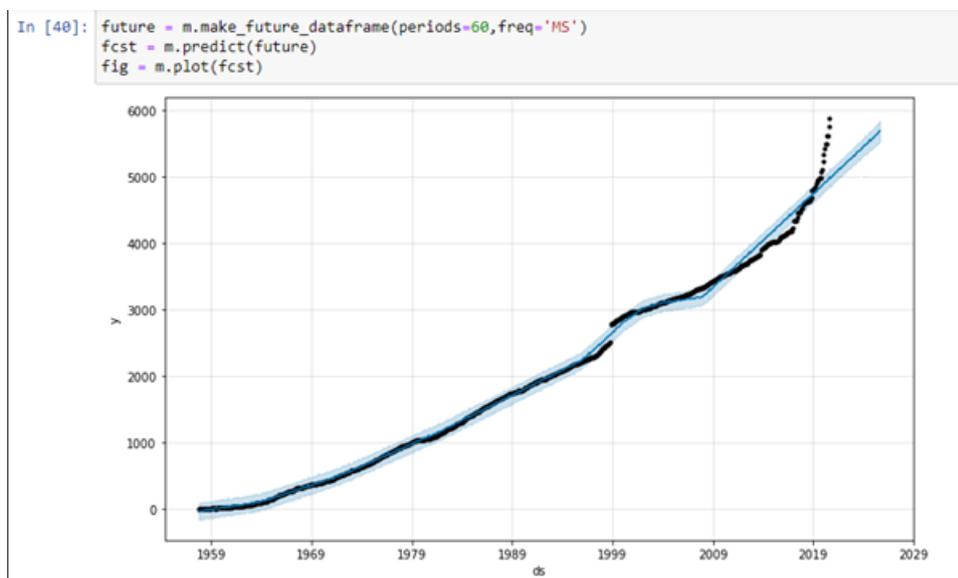


Figura 29: Resultados del modelo de predicción utilizando toda la información disponible. Fuente: Elaboración propia

Pero si se limita el set de datos de entrada en el modelo de predicción; se muestra un pronóstico mucho más alineado con la tendencia actual. Como se presenta en la Figura 30.

```
In [43]: future = m.make_future_dataframe(periods=12,freq='MS')
fcst = m.predict(future)
fig = m.plot(fcst)
```

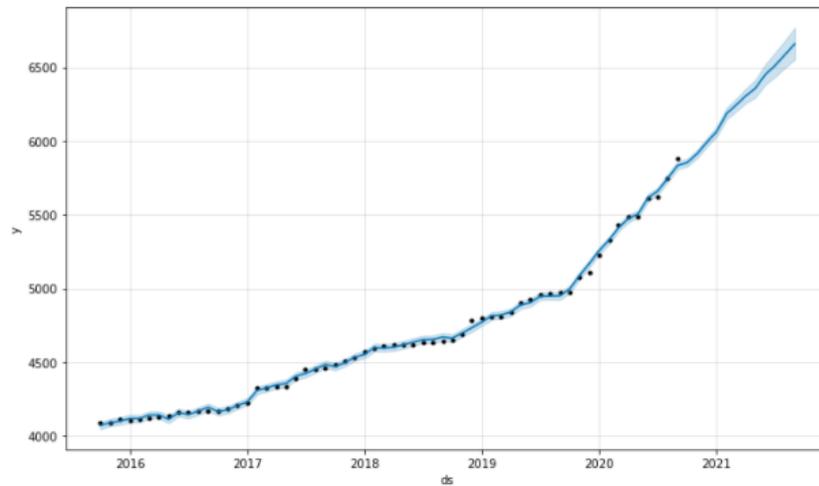


Figura 30: Resultados del modelo de predicción utilizando información de los últimos 5 años. Fuente: Elaboración propia.

Estas predicciones no solo se utilizan como base para mantener una estadística en cuanto a los satélites que orbitan, sino que son insumo para determinar el crecimiento en los tránsitos, como se expone en la sección 5.4.2.

Lo explicado en esta sección se resume en el código python mostrado en la Figura 31 (para mantener el código lo más compacto posible se omitieron las declaraciones de las librerías a utilizar).

```

inputfile = 'SatStatistics.csv'
NumMonths=12
Outputfile = 'SatPredictionPH.csv'
data = pd.read_csv(inputfile,parse_dates=['MesAnno'])
data['ds'] = data['MesAnno']
data['y'] = data['NumSats']
data.drop(labels=['MesAnno', 'NumSats'],axis=1,inplace=True)
data =data.head(NumMonthsBack)

m = Prophet()
m.fit(data)

future = m.make_future_dataframe( periods=int(NumMonths),freq='MS')
fcst = m.predict(future)

result = fcst[~fcst.ds.isin(data.ds)]
result.drop(result.columns.difference(['ds','trend']), 1,
inplace=True)
result = pd.merge(data,
                    result[['ds', 'trend']],
                    on='ds',
                    how='outer')
result.sort_values(by=['ds'],inplace=True)
result.columns = ['MesAnno', 'NumSats','Prediction']
result.to_csv(outputfile,index=False)

```

Figura 31: Código utilizado para creación del modelo de predicción. Fuente: Elaboración propia

5.3.9 Exactitud del modelo de predicción

Esta predicción está hecha basada en un horizonte. Este horizonte es el número de días en el futuro en el cual se realiza el pronóstico, típicamente 30, 90, 180 o 365 días en la mayoría de los casos. Por tanto, para cada observación es necesario determinar cuál es el grado de error en el que se incurre en la creación del modelo.

Para un modelo de series de tiempo la validación de los métodos no se puede realizar mediante una validación cruzada “tradicional”. Este, consistiría en tomar datos al azar y separar la muestra en datos de entrenamiento y datos de validación, generalmente dividiendo el conjunto en 70% datos de entrenamiento y 30% en datos de validación para así obtener un cálculo del margen de error obtenido. Esto simplemente porque las observaciones no son intercambiables entre periodos.

Por este motivo, el método utilizado para hacer la validación del modelo de serie de tiempo es utilizando diferentes puntos en el tiempo seleccionados de tal manera que los “horizontes” seleccionados caigan dentro de la historia disponible hasta ese momento y por tanto puedan ser evaluados contra los valores reales. La ventaja de utilizar estas fechas de control en vez de simular y calcular cada fecha independientemente es que economiza recursos computacionales.

Prophet ya cuenta con métodos que permiten hacer esta validación sencilla, de manera que no hace falta implementar nueva lógica para hacer un diagnóstico de los resultados.

Para el caso de la presente investigación, el modelo va a utilizar solamente los últimos 5 años de historia para calcular el siguiente año en cuestión. Por lo que para la evaluación del modelo se utilizaron 3 años para entrenamiento y 2 años para validar el modelo. El código utilizado se puede observar en la Figura 32.

```
from fbprophet.diagnostics import performance_metrics
from fbprophet.diagnostics import cross_validation
df_cv = cross_validation(m, initial='1060 days', period='365
days' , horizon = '365 days')
df_p = performance_metrics(df_cv)
df_p.head(50)
```

Figura 32: Código de evaluación del modelo de predicción. Fuente: Elaboración propia.

Los resultados obtenidos se describen en la Tabla 24. En esta se muestra el valor MAPE (error medio de porcentaje absoluto) que crece desde un 0.3% el primer mes de validación hasta un 8% luego de un año. Lo cual indica que el error o desviación media estimada es del alrededor de un 8% al cabo de un año para los datos obtenidos.

Horizonte	MAPE
29 días	0.4%
60 días	1.6%
90 días	2.2%
121 días	3.5%
152 días	4.4%
180 días	4.4%
181 días	4.5%
211 días	4.7%
212 días	5.1%
241 días	5.7%
242 días	5.7%
272 días	6.1%
273 días	6.7%
302 días	6.8%
303 días	6.6%
333 días	6.6%
334 días	7.5%
364 días	7.5%
365 días	8.3%

Tabla 24: Resultados del modelo de diagnóstico. Fuente: Elaboración propia.

5.4 Presentación y análisis de resultados

Luego de haber realizado la clasificación de los tránsitos, es indispensable elaborar una interfaz que facilite la comprensión y la visualización de los datos. Con el fin de para poder presentarlos de manera más entendible y poder notar con mayor facilidad los momentos de mayor afectación visual durante el día.

El gráfico principal muestra en el eje X el tiempo y en el eje Y la cantidad promedio, adicionalmente, la clasificación del impacto está descrita en la leyenda de este. De esta forma, se muestran los satélites que están visibles en el espacio cada

hora o cada 5 minutos, clasificados por su impacto. Este gráfico permite ser filtrado por fecha, ciudad, clasificación de impacto y por hora del evento.

Adicionalmente, se agrega un indicador de cantidad de satélites activos totales y otra gráfica que muestra el incremento anual de satélites activos.

Como se puede apreciar en la Figura 33, por ejemplo, se tiene la fecha filtrando los días del 19 de noviembre del 2020 al 22 de noviembre del 2020, esto le indica al reporte que tome las cantidades de tránsitos que sucedieron entre esos días y genere un promedio general, se filtran las exposiciones de 1 segundo lo que indica al reporte que utilice el cálculo generado para este tipo de observaciones y se selecciona la ciudad de San José para tomar en cuenta solamente los tránsitos que fueron visibles desde esa ciudad. Para el impacto se tienen seleccionados todos menos el de impacto Nulo, con lo que se filtra del reporte todos los tránsitos que no causaron ningún impacto en las observaciones. Por último, se seleccionan todas las horas que se desean visualizar en el reporte.

De esta forma la visualización mostrará el promedio de tránsitos desglosados por hora o cada 5 minutos. Si se observa el ejemplo, en promedio a las 4 de la mañana suceden 207 tránsitos de impacto bajo y 101 tránsitos de impacto medio para un total de 308 tránsitos.

Se debe considerar que un mismo tránsito puede pasar de clasificación de impacto en el tiempo, es decir, que un satélite que transita a las 04:05 am con un impacto bajo, puede, por la variación de los factores, a las 04:10 am pasar a impacto medio, alto o nulo.

De manera que, la interrupción que se observa en la serie de tiempo (Figura 33) desde las 6 horas hasta las 16 horas, para las categorías “Bajo”, “Medio” y “Alto” está dada por las características del tránsito que, al ser diurno, se reclasifica como “No Relevante” (nulo en el reporte).

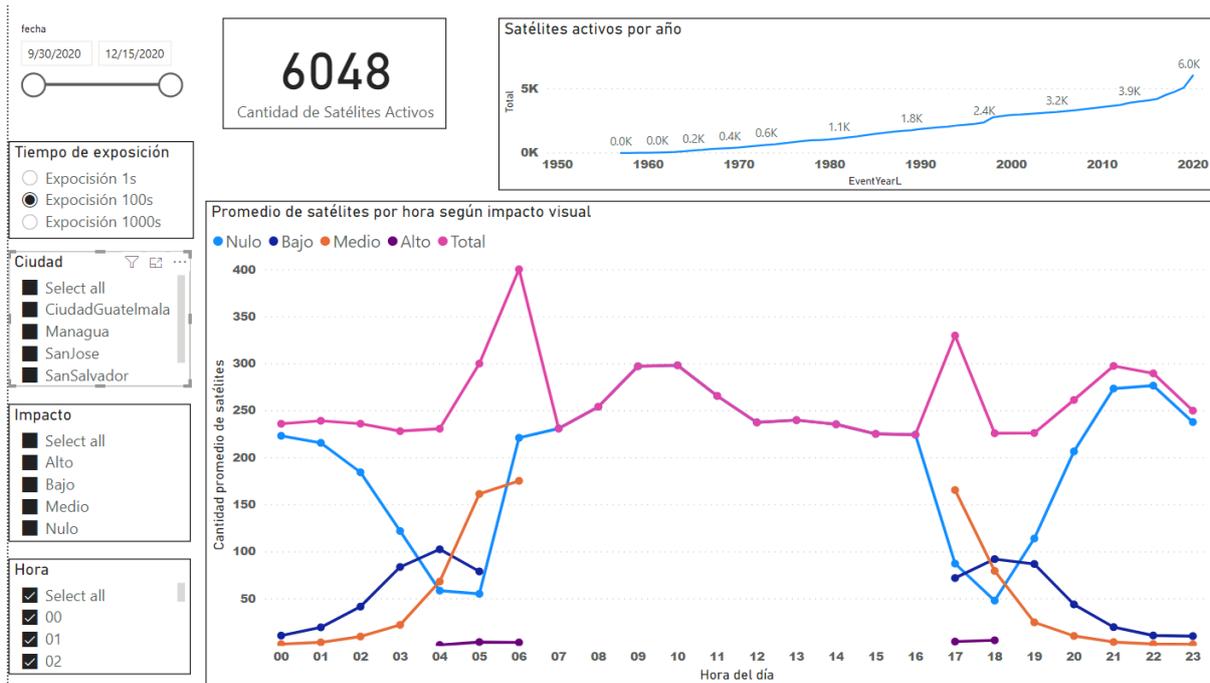


Figura 33: Visualización de clasificación de impacto visual. Fuente: Elaboración propia

5.4.1 Detalle de la interfaz

Fecha: mediante esta casilla se puede establecer el periodo que se desea consultar. Se debe tomar en cuenta que los datos se presentan de forma promediada, por lo que, si en esta casilla se selecciona un rango mayor al día, los datos presentaran el promedio de los días consultados distribuido por hora.



Figura 34: Filtro de fecha. Fuente: Elaboración propia

Tiempo de exposición: se utiliza para seleccionar el tiempo que tarda la fotografía astronómica, según los valores establecidos por Hainaut y Williams (2020). En donde se describen tres escenarios que son facilitados por la interfaz.



Tiempo de Exposición

- Exposición 1s
- Exposición 100s
- Exposición 1000s

Figura 35: Filtro de tiempo de exposición. Fuente: Elaboración propia

Ciudad: muestra un catálogo de las capitales de Centroamérica, que corresponden a la población meta.

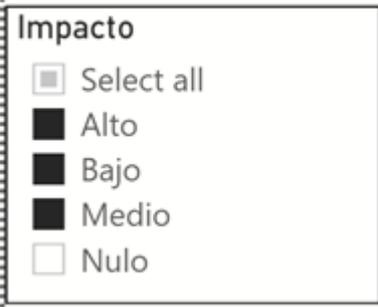


Ciudad

- Select all
- CiudadGuatemala
- Managua

Figura 36: Filtro de ciudad. Fuente: Elaboración propia

Impacto: desde esta casilla, se podrá filtrar el nivel de impacto que poseen los satélites ajustándose a las necesidades de los usuarios.

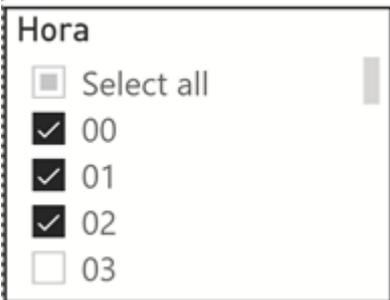


Impacto

- Select all
- Alto
- Bajo
- Medio
- Nulo

Figura 37: Filtro de impacto. Fuente: Elaboración propia

Hora: mediante esta casilla los usuarios pueden escoger el horario que mejor se ajuste a sus intereses.



Hora

- Select all
- 00
- 01
- 02
- 03

Figura 38: Filtro de hora Fuente: Elaboración propia

Cantidad de satélites activos: muestra la cantidad total de satélites activos y en órbita. Este número va a depender de los lanzamientos que se generen y se debe tomar en cuenta que, no necesariamente, todos impactan en las observaciones simultáneamente.



Figura 39: Tarjeta de cantidad de satélites activos Fuente: Elaboración propia

Satélites activos por año: permite identificar el aumento progresivo de los satélites por año, facilitando su análisis posterior.

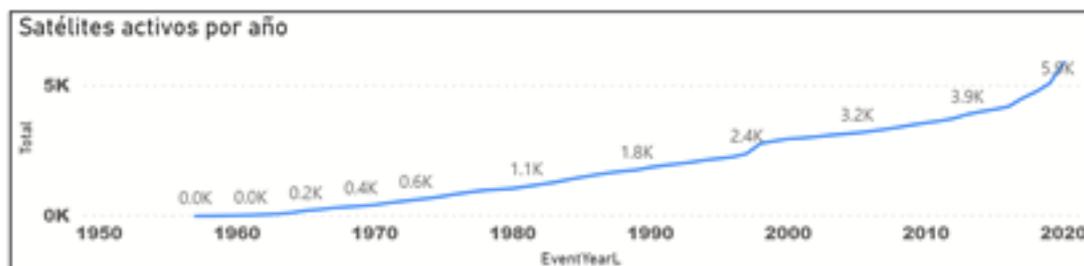


Figura 40: Visualización de crecimiento anual de satélites activos Fuente: Elaboración propia

5.4.2 Modelo futuro

En esta visualización, se utiliza los datos obtenidos del modelo de regresión, para mostrar la predicción de la cantidad de satélites por mes para el siguiente año, basado en el comportamiento de lanzamientos y satélites dados de baja de los años anteriores.

Mediante este modelo predictivo, se busca generar una noción del posible escenario que se presentará los siguientes 12 meses, con lo que los usuarios pueden organizar sus observaciones basados en los datos.

Para esta visualización, se agrega un filtro en donde se puede elegir el año de inicio a partir del cual se muestran los datos, para comodidad del usuario. En donde su eje X presenta los meses y el eje Y la cantidad de satélites, como se puede observar en la Figura 41.

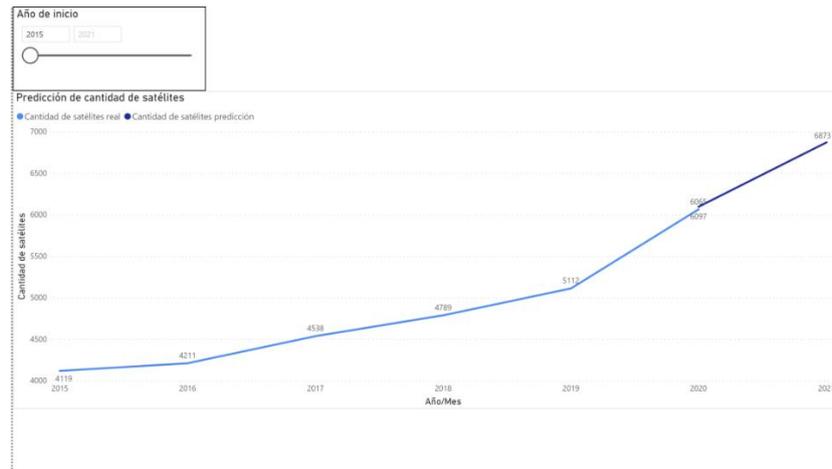


Figura 41: Predicción de cantidad de satélites para el próximo año. Fuente: Elaboración propia

5.4.3 Predicción de tránsitos

Esta visualización muestra la predicción de la cantidad de tránsitos proyectados un año a futuro y clasificados según su impacto, dicho cálculo se realiza basándose en el modelo anterior con el cual se logra obtener un porcentaje de incremento de la cantidad de satélites en un año, luego se realiza una distribución de ese porcentaje sobre la cantidad de tránsitos del último mes de datos reales.

Esto permite comparar el escenario actual y el escenario de predicción a un año, como se muestra en la Figura 42Figura 41.



Figura 42: Predicción de cantidad de satélites para el próximo año. Fuente: Elaboración propia

5.4.4 Análisis de Resultados

El presente estudio evidencia, con datos reales, que el cielo nocturno está siendo continuamente transitado por satélites artificiales y que dichos tránsitos tienen eventualmente, el potencial de afectar las observaciones astronómicas dependiendo de factores como ángulo de elevación del sol, ángulo de elevación del satélite, magnitud aparente del satélite y tiempo de exposición de la observación entre los más importantes.

Los datos de las horas indicadas a continuación se han regionalizado específicamente al territorio Centroamericano por lo que las mismas corresponden a la zona horaria UTC -6.

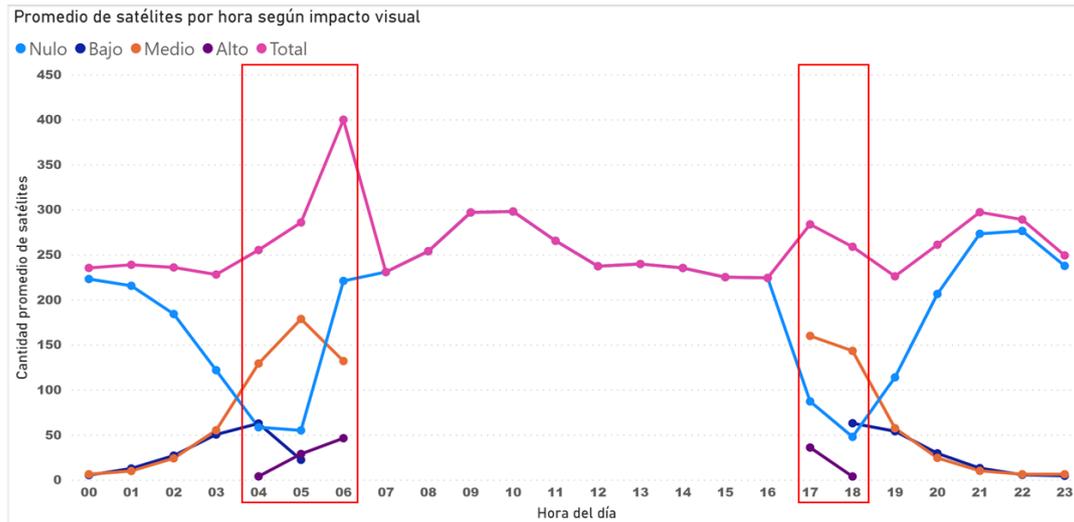


Figura 43: Grafico de clasificación de tránsitos por hora utilizando la configuración de 1000 segundos de exposición.
Fuente: Elaboración propia

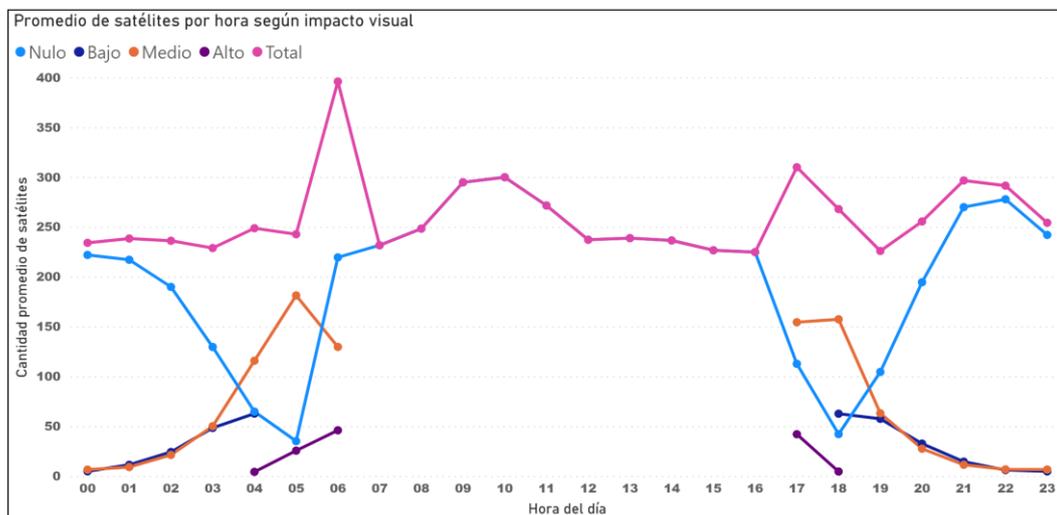


Figura 44: Grafico de clasificación de tránsitos por hora utilizando la configuración de 1000 segundos de exposición para El Salvador. Fuente: Elaboración propia

En la Figura 43 se logra evidenciar que las horas de mayor afectación según los datos consolidados de toda la región de Centroamérica, se dan poco antes del amanecer entre 4 am y 6 am y luego del atardecer entre 5 pm y 6:30 pm. Esto debido a que son las horas en que el sol se encuentra en los ángulos que mayor incidencia tienen sobre los satélites de órbita baja.

También, se determina que entre 9pm y 3am es el mejor momento para realizar observaciones de mayor tiempo de exposición, pues hay menor riesgo de tener imágenes afectadas por el brillo reflejado por los satélites.

Adicionalmente, se observa en la Figura 44 que el escenario es similar si se filtran los datos para un solo país, en este caso, El Salvador, por lo que se puede hablar de una afectación similar en todo el territorio.

Esta información permite programar las observaciones en momentos donde exista menor probabilidad de afectación visual, lo que implica un ahorro de recursos, pues se aumenta la efectividad de cada imagen capturada.

Por medio de la predicción de la cantidad de satélites en órbita un año a futuro, utilizando un algoritmo de series de tiempo, podemos observar que con el transcurso del tiempo la cantidad de satélites orbitando la tierra va en aumento, lo que significa por consiguiente un aumento en la cantidad de tránsitos.

Según el análisis realizado, el crecimiento interanual de satélites activos desde 1999 hasta 2016 era de 2.3%. Este crecimiento desde 2016 hasta 2020 fue del 10% con tendencia al alza. Actualmente existen un poco más de seis mil satélites activos y en un año se pronostica un aumento de alrededor de 800 satélites. Esto indica, como se observa en la Figura 45, un aumento entre 12% y 13%. Significa también, un aumento en la cantidad de tránsitos en un porcentaje similar, como se puede observar en la Figura 46.

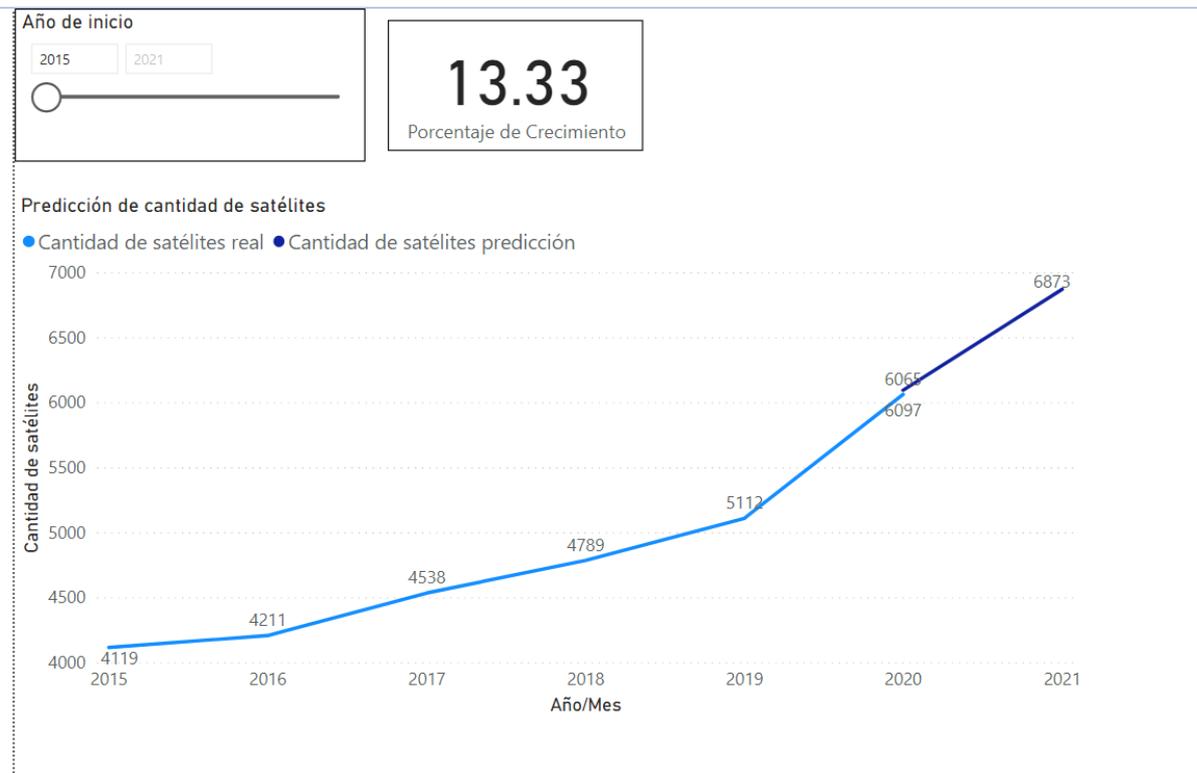


Figura 45: Visualización de predicción del crecimiento en la cantidad de satélites para el año 2021. Fuente: Elaboración propia.



Figura 46: Cálculo de aumento de tránsitos basado en porcentaje de aumento pronosticado. Fuente: Elaboración propia.

Sin embargo, este pronóstico es un poco conservador, considerando que, si comparamos el porcentaje de crecimiento anual de años anteriores, la diferencia entre un año y el siguiente es cada vez más marcada. Debemos recordar que los valores bajos de aumento en la cantidad de satélites en los primeros años surten un efecto compensatorio sobre el mayor incremento que ha habido en los años recientes.

Conociendo que las tecnologías espaciales cada vez van siendo más accesibles al mercado común, es muy probable que el aumento de la cantidad de satélites orbitando la tierra año con año sea mucho mayor, lo que significa un incremento en el impacto visual en las observaciones astronómicas mayor a lo esperado, si no se toman acciones prontas que regulen esta actividad.

5.5 Instrucciones de Implementación del sistema

Los scripts y archivos necesarios para esta implementación han sido puestos a disposición en un repositorio de github, el cual puede ser accedido de manera pública siguiendo el enlace: <https://github.com/oscarcampos-c/TFG-SAT/releases/tag/v1.0> Los contenidos de la carpeta “Implementación” deben ser descargados a una máquina local o virtual.

Como prerrequisito, la persona ejecutando estos scripts debe tener accesos de administrador en la suscripción de Azure donde se vayan a crear los recursos.

Todas las operaciones ejecutadas de manera automática en la implementación se pueden realizar manualmente desde el portal de Azure, pero para hacer la implementación más sencilla se dispuso a utilizar scripts automáticos para la creación de los mismos.

Estos scripts deben ser ejecutados en un ambiente Windows mediante la consola de Powershell. Para esto debe instalarse la librería de Azure para PowerShell utilizando el comando:

```
Install-Module -Name Az -AllowClobber -Scope CurrentUser
```

Luego de ejecutar el comando, debe aceptar instalar todas las dependencias del sistema, cuando se le pida confirmación. Este procedimiento se observa en la Figura 47.

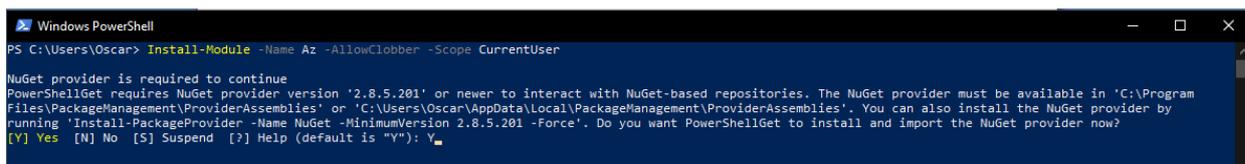


Figura 47: Ejemplo de comando de instalación de dependencias de azure en PowerShell. Fuente: Elaboración propia

Una vez que ya está la librería debe accederse al script llamado “Recursos.ps1”. Para esto, puede abrir la carpeta donde se descargaron los archivos de implementación y utilizando en combinación la tecla “shift” + clic derecho en un espacio en blanco, puede seleccionar la opción: Abrir ventana de PowerShell aquí. Como se muestra en la Figura 48.

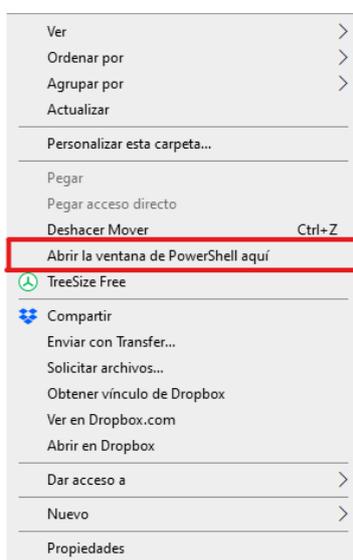


Figura 48: Menú contextual de carpeta en Windows. Fuente: Elaboración propia

Este script de PowerShell recibe 6 parámetros necesarios para la creación de los recursos. Estos parámetros son:

- \$SQLServerUser: El nombre de usuario de la cuenta de administrador del SQL Server.
- \$SQLServerPass: La contraseña del servidor de SQL a crear.
- \$SubscriptionName: Nombre de la suscripción donde se van a crear los recursos.
- \$UsernameSpaceTrack: Nombre de usuario de spacetrack.
- \$PasswordSpaceTrack: Contraseña de spacetrack.

Para ejecutar el script puede utilizar un comando desde la consola de PowerShell como el mostrado en la Figura 49

```
.\recursos.ps1 -SQLServerUser 'SQLAdmin' -SQLServerPass
'Password1' -SubscriptionName 'Visual Studio Enterprise
Subscription' -UsernameSpaceTrack 'spacetrackuser' -
PasswordSpaceTrack 'spacetrackpassword'
```

Figura 49: Llamada al script de despliegue inicial. Fuente: Elaboración propia

El sistema solicitará las credenciales y luego procede con la creación de los recursos. La Figura 50 muestra un ejemplo de la solicitud de las credenciales de Microsoft.

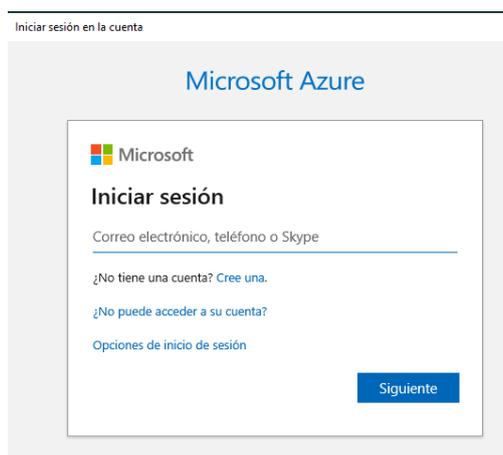


Figura 50: Solicitud de credenciales por parte del script de despliegue. Fuente: Elaboración propia

Una vez que el script se ha ejecutado, los recursos necesarios para la ejecución de los programas se pueden acceder a través del grupo de recursos llamado tfgadf-rg en la suscripción indicada a la hora de la ejecución. Estos están nombrados como nombre del recurso y un Id. Este id es un identificador único creado para evitar que 2 proyectos entren en conflicto ya que el nombre debe ser único a nivel mundial para algunos recursos.

Por ende, los recursos necesarios están nombrados como:

- tfgsatserver<Id>: Servidor SQL Server Azure.
- tfgsat: Base de datos SQL.
- tfgbatch<Id>: Cuenta Batch para procesamiento.
- tfsatstacc<Id>: Cuenta de almacenamiento.
- tfgsatADF<Id>: Recurso de orquestación y movimiento de datos.

Configuración de las aplicaciones

Aunque los recursos sean creados y configurados correctamente con el script, todavía falta subir el código compilado de 4 aplicaciones que son necesarias. El código de todas estas aplicaciones se puede encontrar en el GitHub, ingresando a la carpeta “Código Aplicaciones”. Estas aplicaciones fueron compiladas utilizando la librería pyinstaller y posteriormente comprimidas con el fin de que fueran desplegadas a los sistemas en la nube de manera sencilla cada vez que tuvieran que ser ejecutadas.

Estas aplicaciones son:

- SatEvents: Aplicación basada en Python utilizada para calcular los tránsitos
- FileUploader: Aplicación basada en powershell, miscelánea para subir los archivos resultantes a la cuenta de almacenamiento.
- StandardMag: Programa basado en Python encargado de obtener la magnitud estándar desde Satflare.com
- SatellitePrediction: Igualmente, aplicación basada en Python que devuelve el valor de la predicción.

Para poder subir estas aplicaciones, es necesario abrir desde el navegador el portal de Azure mediante el link <https://portal.azure.com>. Y buscar el grupo de recursos llamado “tfgsat-rg” desde la barra de búsqueda como se muestra en la Figura 51.

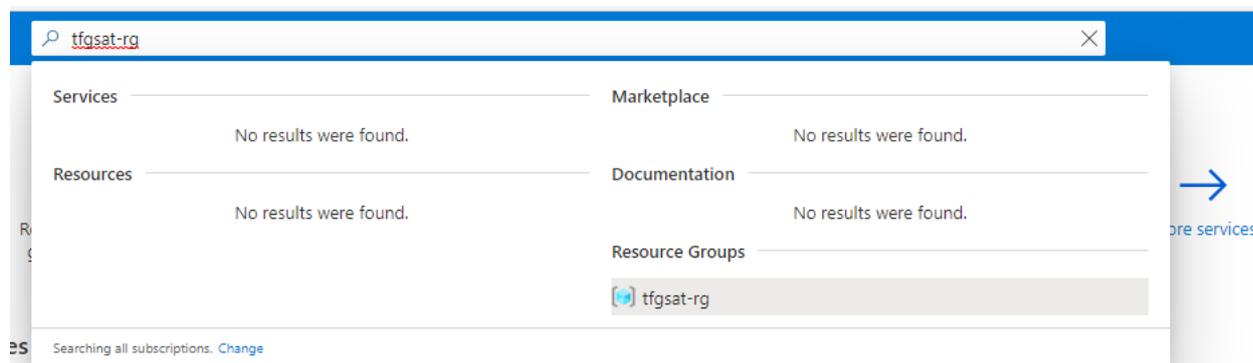


Figura 51: Búsqueda del grupo de recursos implementado. Fuente: Elaboración propia.

En este grupo de recursos debe ser visible una cuenta tipo “Batch”, en la cual se debe ingresar para subir las aplicaciones mencionadas anteriormente. Esta cuenta debe tener un nombre que sigue el formato: tfgbatch<Id> donde el Id

corresponde a una secuencia de números con el fin de hacer que la cuenta sea única por cada implementación. Como se ejemplifica en la Figura 52 resaltado en amarillo.

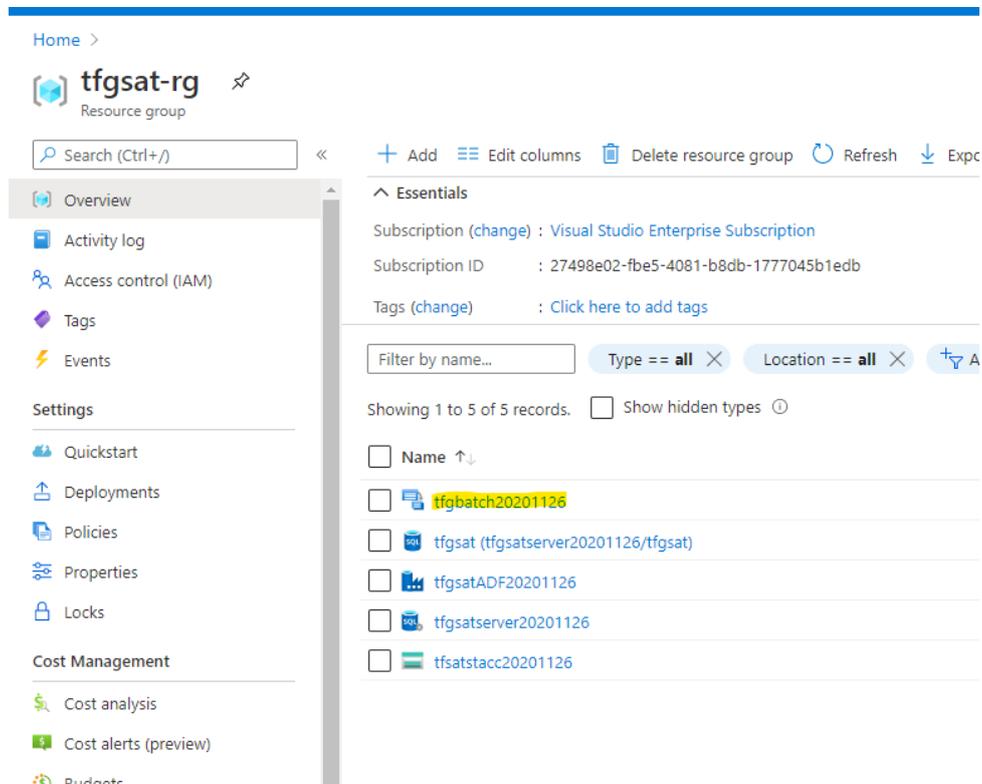


Figura 52: Identificación de la cuenta Batch en el grupo de recursos. Fuente: Elaboración propia.

El siguiente paso es ingresar a la cuenta Batch y seleccionar el menú de aplicaciones. En esta interfaz es donde se deben agregar las 4 aplicaciones, haciendo clic en el botón añadir, siguiendo la Figura 53.

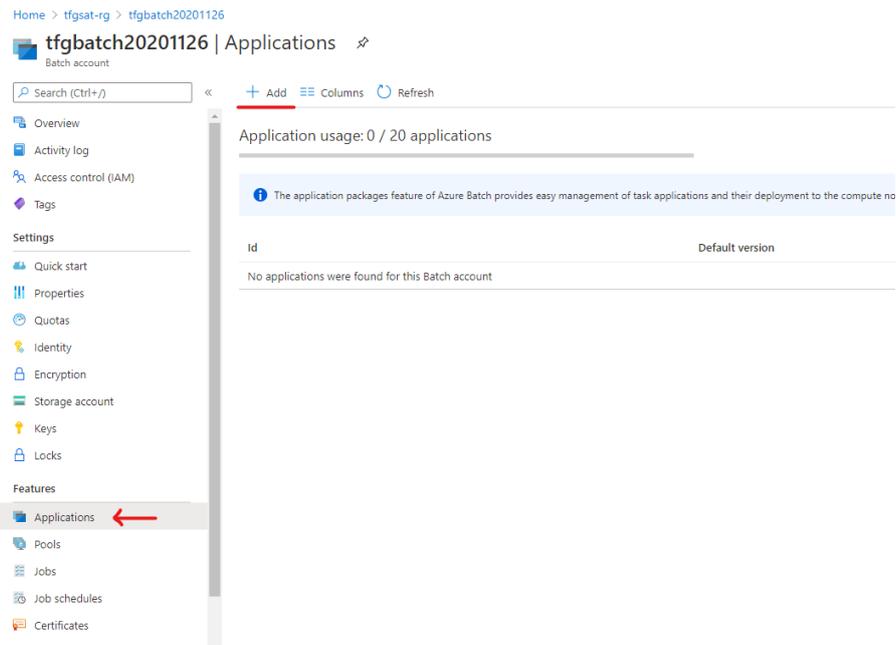


Figura 53: Ejemplo de cómo agregar una aplicación en la cuenta Batch parte1. Fuente: Elaboración propia

Home > tfgsat-rg > tfgbatch20201126 >

New application

Application Id * ⓘ

Version * ⓘ

Application package * ⓘ

Figura 54: Ejemplo de cómo agregar una aplicación en la cuenta Batch parte2. Fuente: Elaboración propia

En Figura 54 se observa la interfaz para agregar cada aplicación. Estas deben tener las siguientes propiedades, en orden:

- Nombre: satevents. Versión: 0.4. Archivo: SatEvents.zip
- Nombre: FileUploader. Versión: 2. Archivo: uploadoutput.zip
- Nombre: StandardMag. Versión: 1. Archivo: GetStandardMag.zip
- Nombre: satelliteprediction: Versión: 2.1. Archivo: SatsFBProphet.zip

En el github que acompaña este proyecto se puede encontrar tanto el código fuente como los archivos comprimidos en formato .zip para ser subidos a la nube de Microsoft.

Por lo que una vez listos, estos deberían lucir como se ejemplifica en Figura 55.

Application usage: 4 / 20 applications

i The application packages feature of Azure Batch provides easy management of task applications and their deployment.

Id	Display name
FileUploader	
SatellitePrediction	
SatEvents	
StandardMag	

Figura 55: Vista final de las cuatro aplicaciones creadas. Fuente: Elaboración propia

Ahora que los códigos binarios de las aplicaciones están correctamente configurados en el sistema, es necesario ligarlas al pool de aplicaciones creado para el fin de ejecutar los binarios. El script ejecutado de despliegue creó el pool necesario, por lo que acá solamente deben enlazarse las aplicaciones.

El primer paso será hacer clic en el pool Windows4cores y una vez seleccionado es necesario seleccionar las cuatro aplicaciones necesarias con su número de versión correspondiente utilizando el control desplegable. Como lo ejemplifican la Figura 56 y la Figura 57.

tfgbatch20201126 | Pools

Batch account

Search (Ctrl+/) << + Add + Add (JSON editor) Columns Refresh

Queries: All pools

Advanced query

Search for pool by ID or display name

Pool ID	Display name	Dedicated nodes	Low-pri
Windows4cores	Windows4cores	0	0

Figura 56: Guía para enlazar las aplicaciones al pool parte 1. Fuente: Elaboración propia.

Windows4cores | Application packages

Pool

Search (Ctrl+/) << + Save X Cancel

Overview

General

Properties

Nodes

Settings

Certificates

Start task

Application packages

Scale

Metadata

If you add application packages to an existing pool, you r

Application	Version
fileuploader	2
satelliteprediction	2.1
satevents	0.4
standardmag	1

Figura 57: Guía para enlazar las aplicaciones al pool parte 2. Fuente: Elaboración propia.

Configuración de la base de datos

En cuanto a la base de datos solamente hay que abrir una nueva sesión de SQL Server en un cliente de base de datos como el SQL Server Management Studio, conectarse con las credenciales utilizadas en los pasos anteriores, seleccionar la base de datos y ejecutando el script “deployment.sql” como se ejemplifica en la Figura 58 y Figura 59.

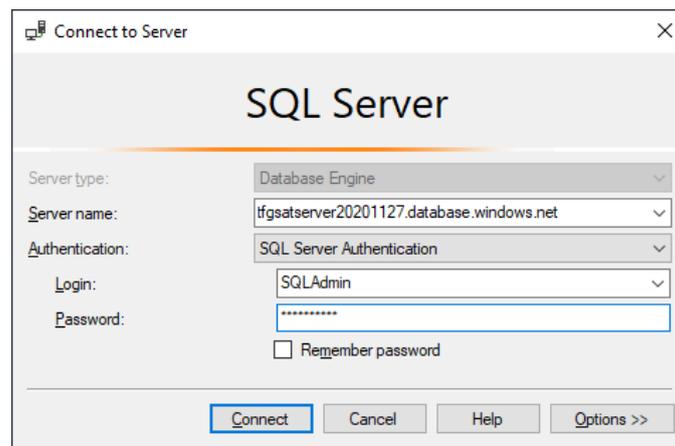


Figura 58: Ejemplo de conexión a base de datos utilizando SSMS. Fuente: Elaboración propia.

Para ejecutar solamente debe abrirse el script con el Management Studio y seleccionar la base de datos de destino como se muestra en verde en la Figura 59.

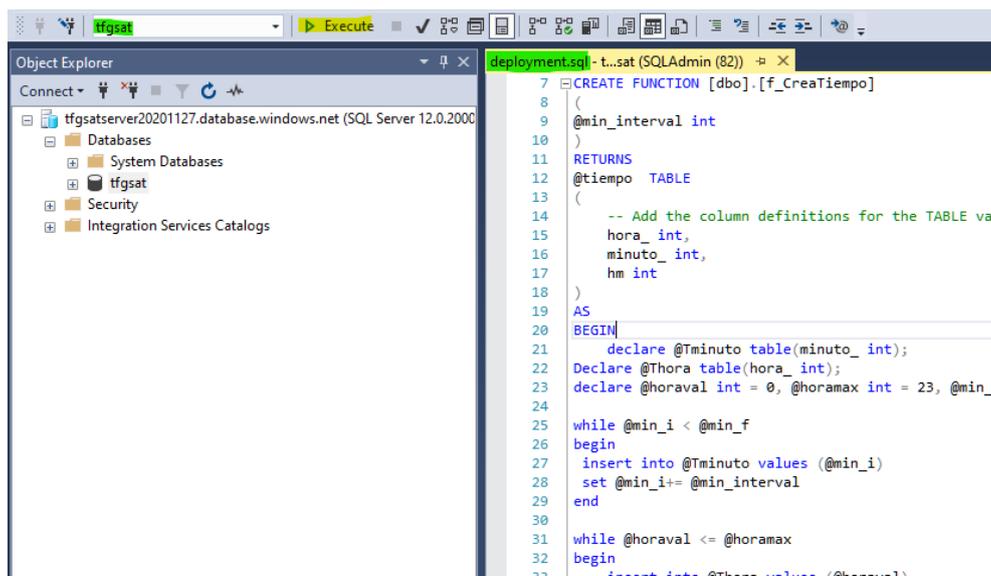


Figura 59: Ejemplo de ejecución de script utilizando SSMS. Fuente: Elaboración propia

Este script se encarga de crear todos los objetos de base de datos necesarios para almacenar y consultar la información. Esta lista se detalla en la Tabla 25

Nombre de objeto	Tipo de objeto	Descripción del contenido
dbo.Cities	Tabla	Información de elevación y coordenadas de las 5 capitales centroamericanas.
dbo.satcat	Tabla	Catálogo de satélites.
dbo.SatEvents	Tabla	Tránsitos satelitales calculados.
dbo.StdMag	Tabla	Magnitud estándar de los satélites.
dbo.stg_satcat	Tabla	Almacena el catálogo de satélites en formato texto de manera temporal.
dbo.stg_stdmag	Tabla	Almacena el conjunto de datos de satflare en formato texto de manera temporal.
dbo.stg_tle	Tabla	Almacena cada lote de información de TLE en formato texto de manera temporal.
dbo.tle	Tabla	Información de los TLE.
dbo.SatPrediction	Tabla	Guarda el resultado de la predicción de datos explicada en la sección 5.3.8 .
dbo.LastSatPrediction	Vista	Obtiene los datos de la predicción más reciente.
dbo.SatCat_Valid	Vista	Catálogo de satélites purgado aplicando las condiciones especificadas en la sección 4.1 .
dbo.StdMag_Calc	Vista	Vista que calcula obtiene los valores de magnitud estándar existentes y aplicando el cálculo de la sección 5.3.6.1 para los que no lo tienen.
dbo.TLE_Latest	Vista	Obtiene el TLE más reciente para cada satélite.

dbo.v_SatEvents	Vista	Devuelve los tránsitos satelitales
dbo.v_SatEventsClas	Vista	Clasifica los tránsitos de acuerdo a lo expuesto en la sección 5.3.7
dbo.v_satEventsPred	Vista	Devuelve la información de predicción necesaria para el reporte.
dbo.GetSatStats	Procedimiento	Calcula la cantidad de satélites activos por mes y año.
dbo.GetTLEIterations	Procedimiento	Devuelve una fila de datos por lote para enviar la solicitud web al servicio de spacetrack.
dbo.InsertTLE	Procedimiento	Mueve los datos desde el almacenamiento temporal en la tabla stg_tle hacia la tabla tle en su formato óptimo de almacenamiento.
dbo.MergeStdMag	Procedimiento	Mueve los datos desde el almacenamiento temporal en la tabla stg_stdmag hacia la tabla StdMag en su formato óptimo de almacenamiento.
dbo.SP_MergeSatcat	Procedimiento	Mueve los datos desde el almacenamiento temporal en la tabla stg_satcat hacia la tabla satcat en su formato óptimo de almacenamiento.
dbo.f_CreaTiempo	Función	Devuelve todas las combinaciones de hora y minutos por cada 5 minutos.

Tabla 25: Objetos de base de datos necesarios almacenar y calcular la información. Fuente: Elaboración propia

Capítulo 6. Conclusiones y Recomendaciones

6.1 Conclusiones

Sobre el objetivo específico: “Adquirir los datos necesarios para el cálculo del tránsito satelital utilizando la base de datos Space-Track de la fuerza aérea de los Estados Unidos”, se concluye:

- La utilización de las interfaces brindadas por Space-Track permiten un acceso ágil y rápido a los datos, de manera que resultó relativamente sencilla la extracción de la información, tanto de satélites como de TLEs.
- Se identificó que alrededor del 6% de los satélites rastreados no poseen información de TLE recientes (últimas 24h desde la hora de consulta).
- La base de datos de satélites provee muy poca información referente a las características físicas de estos. No es posible identificar el tamaño, material o forma de los satélites utilizando los conjuntos de datos facilitados por el servicio.

Sobre el objetivo específico: “Transformar los datos para su correcto almacenamiento en una estructura de base de datos mediante un proceso automatizado de extracción, transformación y carga”, se concluye:

- Dado que todos los datos utilizados son estructurados, no fue necesaria la utilización de una base de datos NoSQL ni hubo que transformar los datos de la fuente, más que solamente agregar campos de control o renombrar algunas columnas para una mayor claridad o por incompatibilidad.
- Fue posible realizar todas las operaciones necesarias utilizando la arquitectura en la nube de Microsoft Azure. Esta plataforma pone a disposición todas las tecnologías necesarias para la creación de los procesos de extracción, procesamiento y almacenamiento de la información.

Sobre el objetivo específico: “Emplear algoritmos de ubicación satelital con el fin de predecir la posición en el cielo del observador”, se concluye:

- La librería Skyfield permitió realizar los cálculos de una manera muy ágil, ya que implementa el algoritmo SGP4/SDP4 para leer los elementos de 2 líneas, además de implementar funciones específicas para determinar la posición del satélite.
- Con el fin de ejecutar el código programa necesario para realizar los cálculos hubo que utilizar una máquina virtual. No fue posible ejecutarlo desde Azure

Data Factory con un componente nativo, aunque esta última herramienta sí permitió integrar este programa en la línea principal de ejecución.

En cuanto al objetivo específico: “Clasificar los satélites de acuerdo a su distancia y magnitud aparente (brillo) promedio para asignar el peso de su posible impacto en las observaciones astronómicas con técnicas de minería de datos.”, se concluye:

- Para lograr la clasificación de los satélites fue necesario utilizar código de SQL que se añadió a la base de datos por medio de una vista, la cual realiza la lectura de los registros de tránsitos y transforma los valores de cada factor determinante a su respectiva categoría.
- Esta clasificación permite mostrar la cantidad de satélites que representan un mayor impacto en las observaciones en un momento determinado, lo que facilita el aprovechamiento de los recursos en las observaciones astronómicas.
- El conocer las horas donde existe mayor impacto visual, los observatorios pueden organizar los momentos de observación donde tengan un menor impacto visual.

Con respecto al objetivo específico: “Confeccionar un reporte que muestre el recuento de los satélites que presentan algún grado de interferencia en la observación y su clasificación en un momento determinado.”, se concluye:

- A través de la herramienta gratuita Power BI Desktop se confecciona el reporte que muestra el grado de interferencia de los satélites clasificándolo en impacto: Nulo, Bajo, Medio y Alto.
- Este reporte permite conocer el nivel de impacto con un detalle de 5 minutos para facilitar su visualización, debido a que, en este rango la variación puede ser significativa, no así en tiempos menores. Lo que además evita la saturación de datos en el reporte.

En referencia al objetivo específico: “Diseñar un reporte que muestre la estimación de la cantidad de satélites proyectado a un año.”, se concluye:

- Mediante la minería de datos, específicamente, con el algoritmo de series de tiempo, se logró estimar la cantidad de satélites activos un año a futuro, lo que puede permitir a los observadores considerar un posible escenario en sus observaciones a futuro.

6.2 Recomendaciones

Luego de obtener y analizar los datos provenientes de las diferentes fuentes de información se plantean las siguientes recomendaciones y acciones estratégicas dirigidas a todos los interesados en observación astronómica sobre la utilización de los datos mencionados.

“Adquirir los datos necesarios para el cálculo del tránsito satelital utilizando la base de datos Space-Track de la fuerza aérea de los Estados Unidos.”

Como primera recomendación es importante utilizar la tecnología final esperada para la obtención y tratamiento de los datos. Esto evita que sea necesario luego tener que migrar el código o la metodología desarrollada a otra arquitectura o lenguaje de programación. De manera que se dedique un tiempo previo a determinar si todo código o herramienta a utilizar se puede integrar con la arquitectura sin tener que hacer erogaciones de dinero no adicionales.

Así también se recomienda extraer una muestra de los datos a realizar y determinar si todos los atributos disponibles son los suficientes para poder responder a todas las cuestiones planteadas. Así, se evitan re trabajos en el momento de aplicar las metodologías ya planeadas.

“Transformar los datos para su correcto almacenamiento en una estructura de base de datos mediante un proceso automatizado de extracción, transformación y carga.

Utilizar una metodología orientada a datos no estructurados y ambientes NoSQL, puede ayudar a que la solución de datos sea más escalable en el futuro, sin embargo, aumenta los costos de mantenimiento con respecto a una base de datos estructurada. Por lo que, si no se deseara recurrir a una solución de BigData, es posible seguir con la tecnología SQL, la cual es más barata (en este contexto), sin embargo, podría no ser muy escalable y presentar problemas de rendimiento en el mediano plazo, principalmente si se desea seguir manteniendo un historial de los eventos generados.

“Emplear algoritmos de ubicación satelital con el fin de predecir la posición en el cielo del observador.”

Dado que el proceso de cálculo es bastante intensivo en cuestión de recursos para las 5 ciudades centroamericanas, es recomendable desde el inicio implementar un procedimiento que haga uso de todos los recursos disponibles de CPU, de otra forma el proceso de carga se vuelve lento. Esto

se puede lograr en Python con la metodología de multiprocesamiento y por ende el diseño del programa se realizaría con esta metodología en mente desde el inicio.

“Clasificar los satélites de acuerdo a su distancia y magnitud aparente (brillo) promedio para asignar el peso de su posible impacto en las observaciones astronómicas con técnicas de ingeniería de datos.”

El proceso de clasificación representa una carga importante para el motor de base de datos, por lo que se recomienda realizar un proceso extra de mantenimiento que limpie datos antiguos que no se requieren para los demás procesos, ya que conforme aumenta el tamaño de la tabla principal del proceso el rendimiento disminuye, causando que la clasificación implique mayor tiempo de ejecución si no se aumenta la capacidad de procesamiento, lo que se traduce en un mayor gasto económico.

“Confeccionar un reporte que muestre el recuento de los satélites que presentan algún grado de interferencia en la observación y su clasificación en un momento determinado.”

Se recomienda valorar el nivel de detalle de tiempo a utilizar en los datos para no saturar los reportes y a la vez mantener la utilidad de la información. Debido a que, si se presentan los datos en rangos de tiempo menores, la variación del impacto puede no ser significativa y causaría ilegibilidad de los datos presentados. De igual forma, si se utilizan rangos muy amplios la información presentada podría perder valor para su análisis.

“Diseñar un reporte que muestre la estimación de la cantidad de satélites proyectado a un año.”

Se considera importante realizar la ampliación de las proyecciones a un mayor nivel de detalle con el que se pueda conocer el escenario en cualquier momento a futuro, no solo en días sino también, en horas y cada cinco minutos. Para ayudar en la gestión de recursos de los observatorios.

Capítulo 7. Reflexiones Finales

La presente investigación surge debido a la incertidumbre del panorama real actual de la región centroamericana con respecto a la afectación visual que tienen los satélites artificiales en las observaciones astronómicas, además impulsados por la preocupación de la comunidad astronómica por la cantidad de satélites que se prevén lanzar dentro de pocos años por varias compañías, se visualizó desarrollar una herramienta que permita monitorear, el impacto actual y proyectado a 12 meses. Esto se logró consumiendo datos oficiales de los sitios www.space-track.com y de www.satflare.com y empleando tecnologías en la nube en conjunto con técnicas de minería de datos.

Al momento de iniciar con esta investigación, descubrimos que existen muy pocos trabajos similares a la fecha, y los existentes realizaban estudios basados en escenarios hipotéticos, lo que nos hace instar a continuar explorando posibilidades de investigación en este campo. Y algunos otros que utilizan datos reales, pero se centran en los grandes observatorios en otras regiones.

Tuvimos la dicha de contar con el apoyo del observatorio Micro Macro de la Universidad Don Bosco de El Salvador, ellos nos aportaron valiosas recomendaciones que enriquecieron nuestro trabajo y nos motivaron a seguir adelante con la investigación, así como las sugerencias para ampliar este esfuerzo en futuras investigaciones.

Por otra parte, hemos aprendido que la utilización de tecnologías en la nube permite desarrollar casi cualquier tipo de proyecto a un bajo costo y permite realizar implementaciones posteriores de toda la infraestructura de manera sencilla por medio de plantillas, lo que facilita enormemente desplegar laboratorios similares en cualquier momento. Además, las tecnologías en la nube permiten escalar recursos a demanda, lo que flexibiliza la plataforma a responder según las necesidades de procesamiento o almacenamiento según se requiera.

Aunado a la flexibilidad de las tecnologías en la nube es importante recalcar, la utilización de librerías de inteligencia artificial como las series de tiempo, lo que nos facilitó la generación de las proyecciones en cantidad de satélites 12 meses a futuro, lo cual nos permitió mostrar un escenario probable del impacto enriqueciendo la investigación.

Por último, un gran desafío que encontramos con el que no contábamos, es la cantidad de datos que se generan por día en cuanto a tránsitos satelitales. Se producen unos 20 mil registros al día por país, lo cual obligó a tener que incrementar el tamaño del clúster utilizado para algunos de los análisis, dado que las consultas eran bastante complejas.

Capítulo 8. Trabajos a Futuro {Líneas Futuras de Investigación}

Durante el transcurso de esta investigación se han observado algunos posibles puntos de mejora o ampliación, los cuales, por temas de alcance, no se incluyeron en esta investigación.

Uno de estos es la ampliación de la predicción de tránsitos a un nivel de detalle de 5 minutos, con lo cual se pueda desarrollar una visualización del promedio de tránsitos y su posible impacto para un momento específico del día.

Las visualizaciones creadas fueron pensadas para uso privado, sin embargo, se puede ampliar el alcance y crear una aplicación pública que permita el acceso de cualquier usuario registrado y generar datos personalizados a una ubicación específica, sin embargo, para esto podría ser necesaria la aprobación por parte de SpaceTrack.org y registrarse como redistribuidor de datos.

El presente trabajo utiliza como base la clasificación desarrollada en el estudio de Hainaut & Williams (2020). Esta clasificación, sin embargo, realiza varias suposiciones que podrían mejorarse o complementarse con datos de otras fuentes, que apoyen a conseguir un valor más detallado o más exacto.

Por último, en la presentación de resultados al observatorio se mencionó el interés sobre la posibilidad de ampliación de este sistema para poder mostrar los tránsitos de manera detallada, especificando una hora, fecha y parte del cielo a observar, con el fin de tener mayor granularidad en la planeación de las observaciones.

Referencias

Referencias bibliográficas

- Barrantes, R. (2008). *Investigación: Un camino al conocimiento*. San José, San José: EUNED.
- Carvalho, R. A., Estela, J., & Langer, M. (2020). *Nanosatellites space and ground technologies, operations and economics*. Hoboken, NJ: Wiley.
- Chen, F., Liu, J., Gou, D., Zhang, X., Chen, L., & Liao, H. (2020). An accurate and universal approach for short-exposure-time microscopy image enhancement.
- Computerized Medical Imaging and Graphics, 83, 101743. doi:10.1016/j.compmedimag.2020.101743
- Cheng, J. (2009). *The Principles of Astronomical Telescope Design*. New York, NY: Springer US.
- Cook, T. D., Reichardt, C. S., & Manuel, A. M. (2005). *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid: Ediciones Morata.
- Dong, W., & Chang-Yin, Z. (2010). An Accuracy Analysis of the SGP4/SDP4 Model. *Chinese Astronomy and Astrophysics*, 34(1), 69-76. doi:10.1016/j.chinastron.2009.12.009
- Gallozzi, S., Scardia, M., & Maris, M. (2020, February 4). Concerns about ground based astronomical observations:A step to Safeguard the Astronomical Sky. *ArXiv*.
- Hainaut, O. R., & Williams, A. P. (2020). Impact of satellite constellations on astronomical observations with ESO telescopes in the visible and infrared domains. *Astronomy & Astrophysics*, 636. doi:10.1051/0004-6361/202037501
- Hejduk, M. (2011). *Specular and Diffuse Components in Spherical Satellite Photometric Modeling*. Retrieved September 1, 2011, from https://www.researchgate.net/publication/253155412_Specular_and_Diffuse_Components_in_Spherical_Satellite_Photometric_Modeling

- Jaén, M. (2020). FEIR 40: Modelos de Regresión. Recuperado de: <http://gauss.inf.um.es/feir/40>
- Liou, J. -, Kieffer, M., Drew, A., & Sweet, A. (2020). The 2019 U.S. Government Orbital Debris Mitigation Standard Practices. *Orbital Debris Quarterly News*, 24(1), 4-9.
- Martin, D. W., & Dávila, J. (2008). *Psicología experimental: Cómo hacer experimentos en psicología*. México, Mexico: Cengage Learning.
- McDowell, J. C. (2020, April 6). The Low Earth Orbit Satellite Population and Impacts of the SpaceX Starlink Constellation. Recuperado de: <https://iopscience.iop.org/article/10.3847/2041-8213/ab8016>
- Maley, P. D., & Pizzicaroli, J. C. (2003). The visual appearance of the Iridium® satellites. *Acta Astronautica*, 52(8), 629-639. doi:10.1016/s0094-5765(02)00127-3
- Kauderer, A. (2011, September). Human Space Flight (HSF) - Realtime Data. Retcuperado de: https://spaceflight.nasa.gov/realdata/sightings/SSapplications/Post/JavaSSOP/SSOP_Help/tle_def.html
- Namakforoosh, M. N. (2005). *Metodología de la investigación* (2nd ed.). Mexico, Mexico: Limusa.
- San-Juan, J. F., Pérez, I., San-Martín, M., & Vergara, E. P. (2017). Hybrid SGP4 orbit propagator. *Acta Astronautica*, 137, 254-260. doi:10.1016/j.actaastro.2017.04.015
- Schaefer, B. E., Barber, M., Brooks, J. J., Deforrest, A., Maley, P. D., Mcleod, N. W., . . . Whitney, S. (1987). The Perseus Flasher and satellite glints. *The Astrophysical Journal*, 320, 398. doi:10.1086/165552
- Taylor, S. J., & Letham, B. (2017). Forecasting at scale. *PeerJ Preprints*. doi:10.7287/peerj.preprints.3190v2
- Thomas Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, pages 199–220, 1993.

- Thompson, R. B., & Thompson, B. F. (2005). *Astronomy hacks*. Beijing: O'Reilly Media.
- Virgili, B. B., Dolado, J., Lewis, H., Radtke, J., Krag, H., Revelin, B., . . . Metz, M. (2016). Risk to space sustainability from large constellations of satellites. *Acta Astronautica*, 126, 154-162. doi:10.1016/j.actaastro.2016.03.034
- Seitzer, P. (2020, June). Mega-Constellations of Satellites and Optical Astronomy. *Bulletin of the American Astronomical Society*, 52(3).
- Wang, R., Liu, J., & Zhang, Q. (2009). Propagation errors analysis of TLE data. *Advances in Space Research*, 43(7), 1065-1069. doi:10.1016/j.asr.2008.11.017
- Wood, L. (2003). Satellite Constellation Networks. *Internetworking and Computing Over Satellite Networks*, 13-34. doi:10.1007/978-1-4615-0431-3_2
- Xiao-Li, X., & Yong-Qing, X. (2019). Study on the Orbit Prediction Errors of Space Objects Based on Historical TLE Data. *Chinese Astronomy and Astrophysics*, 43(4), 563-578. doi:10.1016/j.chinastron.2019.11.007

Referencias electrónicas

- ¿Cuáles son las tarifas por servicios profesionales? (2019, April 29). Recuperado de: <https://www.colegiocienciaseconomicas.cr/faq-items/cuales-son-las-tarifas-por-servicios-profesionales/>
- Adrián, Yirda. (Última edición: 7 de octubre del 2019). Definición de Ingeniería. Recuperado de: <https://conceptodefinicion.de/ingenieria/>
- Agencia Espacial Mexicana. Introducción a la mecánica orbital 3. Recuperado de: http://www.educacionespacial.aem.gob.mx/images_ise/pdf/03mecanica_orbital_3.pdf
- Análisis de regresión - Definición, qué es y concepto | Economipedia. (2020). Recuperado de: <https://economipedia.com/definiciones/analisis-de-regresion.html>
- Astronomía Iniciación (2011). Magnitud aparente y magnitud absoluta. Recuperado de <https://www.astronomia-iniciacion.com/magnitudes.html>

Bartels, M. (2019, June 04). Astronomy Group Says Starlink and Other Satellite Constellations Could Threaten Science. Recuperado en Junio 27, 2020, de <https://www.space.com/astronomy-group-worries-about-starlink-science-interference.html>

Bases de Datos NoSQL | Qué son, marcas, tipos y ventajas. (2020). Recuperado de: <https://www.grapheverywhere.com/bases-de-datos-nosql-marcas-tipos-ventajas/>

Bases de Datos NoSQL: qué son y cuáles son sus ventajas. (2020). Recuperado de: <https://www.unir.net/ingenieria/revista/noticias/bases-de-datos-nosql/549204809132/>

Conceptos de minería de datos. (2020). Recuperado de: <https://docs.microsoft.com/es-es/analysis-services/data-mining/data-mining-concepts?view=asallproducts-allversions>

Dunbar, B. (2015, June 01). What Is an Orbit? Recuperado de: <https://www.nasa.gov/audience/forstudents/5-8/features/nasa-knows/what-is-orbit-58.html>

El Salvador – UDB inaugura el “Observatorio Micro Macro”. (2020). Recuperado de: <https://www.infoans.org/es/secciones/noticias/item/8623-el-salvador-udb-inaugura-el-observatorio-micro-macro>

EUMETSAT. When a satellite activates its thrusters movement takes place - this is known as a manoeuvre. Recuperado de: <https://www.eumetsat.int/website/home/Satellites/LaunchesandOrbits/SatelliteOrbits/Satellitemanoeuvres/index.html>

Honsberg, C., Bowden, S. (2014, August 07). Elevation Angle. Recuperado el 27 de Julio, 2020 de <https://www.pveducation.org/pvcdrom/properties-of-sunlight/elevation-angle>.

IBM Knowledge Center. (2020). Recuperado de: https://www.ibm.com/support/knowledgecenter/es/SS3RA7_sub/modeler_main_help_client_ddita/components/neuralnet/neuralnet_model.html

International Astronomical Union. (2020). Satellite Constellations. Recuperado de: <https://www.iau.org/public/themes/satellite-constellations/>

Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos. (2020). Recuperado de: <https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>

Método de Agrupamiento o Clustering - Ligdi González. (2020). Recuperado de: <https://ligdigonzalez.com/metodo-de-agrupamiento-o-clustering-aprendizaje-no-supervisado/>

Modelado de Datos: Definición, Usos y Tipos. (2020). Recuperado de: <https://www.tecnologias-informacion.com/modeladodatos.html>

Naranjo, L. (2020). *Investigación en Informática: el enfoque alternativo*. CPIC Technology Inside, 5(1 Semestre), 4-14. <https://www.cpic.or.cr/Actas/Download/189>

OMM - Nosotros. (2020). Recuperado el 19 Julio, 2020, Recuperado de: <http://omm.udb.edu.sv/omm/public/pagina/historia>

Resnick, B. (2020, January 07). The night sky is increasingly dystopian. Recuperado el Julio 27, 2021, de <https://www.vox.com/science-and-health/2020/1/7/21003272/space-x-starlink-astronomy-light-pollution>

PowerData, R. (2020). ¿Qué son los procesos ETL? Recuperado de: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/qu-son-los-procesos-etl>

Serrano, G. P. (2010). Modelos de Investigación Cualitativa en Educación Social y Animación Sociocultural. Madrid: Narcea Ediciones.

space-track.org, S. (n.d.). Help Documentation. Recuperado de: <https://www.space-track.org/documentation>

Tom Mitchell's Home Page. (2020). Recuperado de: <http://www.cs.cmu.edu/~tom/>

Www.high-iso.de. Iridium-Flares fotografieren. Recuperado de: <https://www.high-iso.de/blog/tutorials/iridium-flares-fotografieren/>

Xavier Basogain (1998). Recuperado de: https://ocw.ehu.eus/pluginfile.php/9047/mod_resource/content/1/redes_neuro/contenidos/pdf/libro-del-curso.pdf

Apéndices

Apéndice 1:

Carta de aval y presentación de resultados por parte del observatorio Micro Macro de la Universidad Don Bosco:



Ciudadela Don Bosco, 16 de diciembre de 2020.

Señores Universidad Cenfotec

Reciban un cordial saludo de parte del Observatorio Micro Macro de la Universidad Don Bosco (OMM-UDB). A través de la presente, hacemos constar que los ingenieros Oscar Campos Miranda y Guillermo Chinchilla Elizondo, en calidad de estudiantes de maestría, presentaron de manera satisfactoria el sistema y los resultados de su trabajo de graduación denominado: **"Estimación del impacto del tránsito de los satélites en la observación astronómica en el territorio centroamericano utilizando una arquitectura basada en la nube"**, el día Miércoles 16 de diciembre de 2020, a las 8:00 am, de manera virtual. Con este encuentro, el Observatorio Micro Macro visualiza la futura implementación de dicho trabajo para uso interno.

Igualmente, damos la autorización para publicar el nombre del observatorio en el documento final como parte interesada de la investigación.

Atentamente

Brisa Terezón

Directora ejecutiva OMM-UDB