🔭 Title Page

ELISE PROTTI
DR. GARCIA
ASP X
2025-2026

APPLIED SCIENCE PROJECT LOGBOOK

MACHINE-LEARNING ENHANCED ORBIT PROPAGATION: IMPROVING LOW EARTH ORBIT PREDICTION USING TLE AND GPS DATA

📆 Calendar

MONTHLY CALENDAR

See individual months.

Legend:
Light Blue = ASP Class
Hot Pink = Meeting with Mentor
Light Pink = Mentor Meeting Recap Email
Purple = Dr. Garcia Meeting
Red = Due Dates
Yellow = Optional Assignments/Tasks
White = Future Tasks/Tasks in Progress
Light Green = Completed Tasks
Dark Blue = Holidays
Brown= Carry into Next Month
Neon Green = Other
Blue Fill = Weekends

🗓️ September 2025

SEPTEMBER 2025

SUN	MON	TUE	WED	THU	FRI	SAT
	1 ASP CLASS	2 Create mentor meeting agenda	3 ASP CLASS Meeting w mentor at 4:00 pm Send meeting summary + next steps email	4	5 ASP CLASS Research calendar formats (30 mins) FIll in calendar info (20 mins)	6
7	8 Email back mentor confirming Sept. meeting times	9 ASP CLASS Plan Sept. calendar (45 mins) Create mentor meeting agenda (30 mins)	10	11 ASP CLASS Meeting w mentor at 12:30pm Send meeting summary + next steps email (10 mins) Set up Paper-Pile (rest of class, 45 mins)	12 Read research paper guidelines (10 mins)	13 FINISH reading research paper guidelines
14	15 ASP CLASS Meeting w Dr. Garcia (20 mins) Watch "Neural Networks" video + take notes (45 mins) Read DASP Meeting presentation + notes (20 mins) -> search unknown terms	16 Read ESA annual report + notes -> search unknown terms FINISH reading DASP Meeting presentation + notes-> search unknown terms Research types of data + potential data sources Create mentor meeting agenda	17 ASP CLASS Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Research related research articles to find unique project idea (rest of class, 30 mins)	18 Research related research articles to find unique project idea -> look at "further applications" sections of research papers	19 ASP CLASS Meeting w Dr. Garcia (30 mins) Research related research articles to find unique project idea (30 mins) Read Comparative Analysis of Resident Space Object (RSO) Detection Methods" + notes (30 mins) Start Intro for R.P. (30 mins)	20 Plan Oct. calendar Read "Comparative Analysis of RSO Detection Methods" + notes FINISH Comparative Analysis of Resident Space Object (RSO) + notes Research related research articles to find unique

						project idea -> look at "further applications" sections of research papers
21 Set up development environment (Github, Codex, VS)	22 Plan Oct. calendar	23 ASP CLASS Read "A Machine Learning Approach to Space Debris Characterisation and Classification" + notes (45 mins) Work on Intro, for R.P. (45 mins) -> Additional research time interspersed	24 Research into LeoLabs -> what does the company do? What are the commercial applications for the orbital debris and AI field?	25 ASP CLASS Meeting w mentor **after school** (check in about R.P) Send meeting summary + next steps email Draft mentor meeting questions (20 mins) Read "A Machine Learning-Based Approach for Improved Orbit Predictions of LEO Space Debris With Sparse Tracking Data From a Single Station" + take notes Research into LeoLabs	26	27 Plan Oct. calendar + get logbook ready for submission
28 Set up "Cursor AI" in my dev environment (integrate with Codex and Github) Plan Oct. calendar + get logbook ready for submission	29 ASP CLASS LOGBOOK DUE Send email to mentor with updated flowchart + questions about methodology (20 mins) Work on Intro for R.P. (70 mins)	30 Work on Intro for R.P.				

Monthly Goals - September

- Logbook due on Sept. 29
  - Plan out October logbook (calendar)!
- Come up with rough idea of specific project question + methodology
  - Make a flow chart
- Finish points 1-6 on research proposal (rough draft) -> CARRY INTO NEXT MONTH
  - Due Oct. 16, first draft done by Oct. 9 for feedback from mentor -> CARRY INTO NEXT MONTH
- Develop a stronger understanding of:
  - Development environment
  - Types of data (CVS, TLE, image)
  - Current state of AI and orbital debris field of study
- Plan Oct. meeting times

🗓️ October 2025

OCTOBER 2025

SUN	MON	TUE	WED	THU	FRI	SAT
			1 ASP CLASS Extended research into LeoLabs - how commercial opportunities can arise from my project (30 mins) Work on R.P. Title + Intro (45 mins) Read and take notes on the research paper “Improving Orbit Prediction Accuracy…” (45 mins)	2 Create mentor meeting agenda Work on R.P. Title + Intro + research questions Read and annotate research paper “Improving Orbit Prediction Accuracy…”	3 ASP CLASS Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Work on R.P. Intro + research questions (rest of class/45 mins)	4 Set up “Cursor AI” in my dev environment (integrate with Codex and Github)
5	6	7 ASP CLASS Meeting w Ms. Parker (12:50-1:20) Work on R.P. INTRO FIRST DRAFT aim to complete (whole class/90 mins)	8 Work on R.P. Research questions Send R.P. draft to Dr. Garcia	9 ASP CLASS Meeting w Dr. Garcia (20 mins) Work on R.P. research questions + objectives (30 mins) Read and take notes on the research paper “Colliding Satellites: Consequences and Implications” (40 mins)	10 Include R.P. edits from Dr. Garcia + work on R.P. research questions + objectives Read and take notes on the research paper “Colliding Satellites: Consequences and Implications”	11 Include R.P. edits from Dr. Garcia Read and take notes on the research paper “Colliding Satellites: Consequences and Implications”

12	13 THANKSGIVING	14 ASP CLASS Include R.P. edits from Dr. Garcia + finish objectives	15 Create mentor meeting agenda SEND MENTOR PARTIAL DRAFT OF R.P. Include R.P. edits from Dr. Garcia/mentor	16 ASP CLASS Meeting w mentor at 1:00pm Send meeting summary + next steps email (10 mins) Additional research on using Python and pandas to create a scientific dataframe of information (20 mins) Start working on R.P. methodology (25 mins/rest of class)	17 Work on R.P. Methodology	18 Work on R.P. methodology (final draft) + final draft of R.P. w/o significance and final references
19	20 ASP CLASS Finish R.P. significance (30 mins) Finish R.P. correct referencing (10 mins) Edit R.P. methodology + update flowchart to methodology (30 mins) Read and make notes on the research paper "Precise and Efficient Orbit Prediction..." (20 mins)	21 Create mentor meeting agenda Finish final draft of R.P. + send to mentor Read and make notes on the research paper "Orit Precision Analysis of Small Man-Made..." Read and make notes on the research paper "Familiarization with Pandas..."	22 R.P. FINAL DRAFT DUE TONIGHT ASP CLASS Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Edit R.P. with mentor edits by tonight to hand it in (60 mins/rest of class)	23	24 ASP CLASS Start building outline for oral presentation (45 mins) Find presentation template (10 mins) Start working on 1/3 intro slides of OP (35 mins)	25 WEDDING - NO WORK TIME

26	27	28	29	30	31	
WEDDING - NO WORK TIME	Plan November meeting times Plan November calendar	ORAL PRESENTATION DAY 1 ASP CLASS Watch presentations (90 mins) OR Finish 3/3 intro slides of OP (60 mins) Finish 1/1 research questions slide of OP (30 mins) Watch + take notes on "Neural Networks Part 2" Youtube video for enriched understanding (60 mins)	Create mentor meeting agenda Work on OP Methodology and Conclusion slides	ORAL PRESENTATION DAY 2 LOGBOOK DUE ASP CLASS Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Make mentor edits on OP (30 mins) Finish editing + completing OP (30 mins) OR Watch presentations (60 mins)	HALLOWEEN Meeting w Dr. Garcia at lunch to practice OP + get feedback Make mentor edits on OP Finish editing + completing OP	

Monthly Goals - October

- Logbook due on Oct. 30
    - Plan out November logbook (calendar)!
- Oral presentations about R.P. on Oct 28, 30, **November 3rd**
- Research proposal due on Oct. 22
    - Aim to have this finished the first week of October
- Develop a greater understanding of how TLE and GPS data is gathered, how to properly put into into a dataframe (using pandas + Python), and propagate data
- Plan Nov. meeting times

🗓️ November 2025

NOVEMBER 2025

SUN	MON	TUE	WED	THU	FRI	SAT
						1
2	3 **ORAL PRESENTATION DAY 3 (MY PRESENTATION DAY)** ASP CLASS Do my presentation and watch other students' presentations (90 mins)	4 Make mentor meeting agenda	5 ASP CLASS Meeting w mentor at 9:00am Send meeting summary + next steps email (10 mins) Start data acquisition of TLE data (50 mins) for CASSIOPE Complete CYSF portal information (10 minutes)	6 Complete CYSF portal information (if not finished)	7 **ORAL PRESENTATION DAY 4** ASP CLASS Watch other students' presentations (90 mins) OR Data acquisition of GPS truth data at the epochs of study/truth data ephemeris (60 mins) for CASSIOPE Start research into what weather data we will be using for this experiment -> start finding sources that contain the space weather variables for download (30 mins) for CASSIOPE	8
9	10 REMEMBRANCE DAY LONG WEEKEND	11 REMEMBRANCE DAY LONG WEEKEND	12 Continue research into what weather data we will be using for this experiment -> start finding sources that	13 ASP CLASS Meeting w mentor at 12:30pm Send meeting summary + next steps email (10 mins)	14	15

			contain the space weather variables for download Make mentor meeting agenda	Finish downloading LASP (LISIRD) Solar Radio Flux at 10.7cm data AND Kp index data -> all links in logbook daily notes (90 mins)		
16	17 ASP CLASS Research methods to propagate orbits using an SGP4 system in Python (60 mins) General research on using Pandas in Python for data analysis (30 mins)	18 Make mentor meeting agenda	19 ASP CLASS Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Read articles about using Pandas in Python + take notes CONT. (30 mins) Practice implementing sample data sets into Python to learn how to create DataFrames - data from Cisco online course (60 mins)	20	21 ASP CLASS Start inputting our downloaded project data into Python DataFrame (60 mins) Background research about how to use SPG4 integrated into Python - CONT. (30 mins)	22
23	24	25 ASP CLASS MISS CLASS BECAUSE OF PRESIDENT'S BREAKFAST Research methods to calculate error residuals of the collected data (30 mins) Start calculating residuals process	26 Plan December meeting times Finalize logbook for November + make December calendar	27 LOGBOOK DUE ASP CLASS MISS CLASS BECAUSE OF VBALL PROVINCIALS Research methods to calculate error residuals of the collected data (30 mins)	28 Finalize logbook for November + make December calendar Watch videos about using Pandas + NumPy in Python + take notes	29

		(60 minutes)		Start calculating residuals process (30 minutes) Finalize logbook for November + make December calendar (30 mins)		
30						

<u>Monthly Goals - November</u>

- Logbook due on Nov. 27
    - Plan out December logbook (calendar)!
- Complete data acquisition (TLEs, GPS, space weather, residuals)
    - Put all the info in a Python ephemeris using Pandas (if time -> if not, push this to December before Winter Break)
- Develop a greater understanding of how TLE and GPS data is gathered, how to properly put into into a dataframe (using pandas + Python), and propagating data
- Plan Dec. meeting times

🗓️ December 2025

DECEMBER 2025

SUN	MON	TUE	WED	THU	FRI	SAT
	1 ASP CLASS MISS CLASS BECAUSE OF ASSEMBLY	2 Make mentor meeting agenda	3 ASP CLASS Meeting w mentor at 8:30am Send meeting summary + next steps email (10 mins) Meeting w Dr. Garcia (10 mins) Continue working on inputting downloaded data into Python DataFrame (60 mins/any extra time)	4	5 ASP CLASS Finish working on inputting downloaded data into Python DataFrame + making clear headers/titles (60 mins) Make a new document to work on the introduction section paper -> copy and past current intro from R.P. and make edits according to Dr. Garcia edits (15 mins) Research + take notes more about TensorFlow and PyTorch and their differences - "TensorFlow: A system for large-scale machine learning" (15 mins)	6 Finalize incorporating Dr. Garcia edits in intro section + start adding research questions/objectives paragraph Finish research + take notes more about TensorFlow and PyTorch and their differences - "TensorFlow: A system for large-scale machine learning"
7 Finalize incorporating Dr. Garcia edits in intro section + start adding research questions/objectives paragraph Research +	8	9 ASP CLASS Meeting w Dr. Garcia (10 mins) Propagate the orbit of 1 TLE then send this TLE to my mentor to compare accuracy (60 mins) -> Encountered a	10 Make mentor meeting agenda Continue calculating residuals process Send TLE to mentor (if not done so already)	11 ASP CLASS Meeting w mentor at 9:00am Send meeting summary + next steps email (10 mins) Fix TLE propagation by converting to	12 Start calculating residuals and saving the information into a spreadsheet/CSV file -> convert onto the Python file later (30 mins)	13 Finish extended research into TensorFlow + how to build the input dataset - "Analysis of the Application

take notes more about TensorFlow and PyTorch and their differences - "TensorFlow : A system for large-scale machine learning"		problem in the propagations (the error is very high -> different coordinate system?)		ITRF using Python on VS Code (30 mins) Extended research into TensorFlow + how to build the input dataset - "Analysis of the Application Efficiency of TensorFlow and PyTorch in Convolutional Neural Network" (30 minutes/rest of time)		Efficiency of TensorFlow and PyTorch in Convolution al Neural Network"
14	15 ASP CLASS **INTRODUCTION SECTION PAPER DUE TODAY** Meeting w Dr. Garcia (10 mins) Start building the specific ML-ready dataset and save in Python -> export file - might need mentor's help with this (90 mins)	16 Make mentor meeting agenda	17 ASP CLASS Meeting w mentor at 12:30pm Send meeting summary + next steps email (10 mins) Finalize logbook for December + make January calendar (60 mins/rest of time)	18	19 ASP CLASS Meeting w Dr. Garcia (10 mins) Finalize logbook for December + make January calendar (45 mins) Plan February calendar (45 mins) Study for physics unit test (remaining time, if needed)	20 WINTER BREAK
21 WINTER BREAK	22 WINTER BREAK	23 WINTER BREAK	24 WINTER BREAK	25 WINTER BREAK	26 WINTER BREAK	27 WINTER BREAK
28 WINTER BREAK	29 WINTER BREAK	30 WINTER BREAK	31 WINTER BREAK			

Monthly Goals - December

- Logbook due on Jan. 30
  - Plan out January logbook (calendar) after midterms!
- Finish SGP4 propagation stage in December (putting all data into a Python DataFrame, implementing SGP4 algorithm, calculating residuals given the propagated states)
- Start the machine learning process in TensorFlow in December during Winter Break (optional)
- Develop a greater understanding of how TensorFlow works -> specifically what type of format the input data set is needed in, and how to get my dataset to follow these guidelines/data types
- Plan Jan. meeting times

🗓️ January 2026

JANUARY 2026

SUN	MON	TUE	WED	THU	FRI	SAT
				1 WINTER BREAK	2 WINTER BREAK	3 WINTER BREAK
4 WINTER BREAK	5	6 ASP CLASS STUDYING FOR MIDTERMS	7 Make mentor meeting agenda	8 ASP CLASS Meeting w mentor at 1:00pm Send meeting summary + next steps email (10 mins) STUDYING FOR MIDTERMS	9 MIDTERM EXAM BREAK	10 MIDTERM EXAM BREAK
11 MIDTERM EXAM BREAK	12 MIDTERM EXAM BREAK	13 MIDTERM EXAM BREAK	14 MIDTERM EXAM BREAK	15 MIDTERM EXAM BREAK	16 MIDTERM EXAM BREAK	17 MIDTERM EXAM BREAK
18 MIDTERM EXAM BREAK	19 MIDTERM EXAM BREAK	20 ASP CLASS Meeting w Dr. Garcia (10 mins) Continue building ML-ready ephemeris for TensorFlow -> merge propagations ephemeris with space weather ephemeris (60 mins) Fix propagation inaccuracies by adding a leap second to the UTC time (30 mins/rest of class)	21 Make mentor meeting agenda	22 ASP CLASS Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Clean space weather data (Kp and F10.7) by sorting the CSV files for epoch in ITRF and resulting space weather value (45 mins) Finish building ML-ready ephemeris for TensorFlow (45	23 Finish building ML-ready ephemeris for TensorFlow (if not finished during class)	24 Start/finish research + take notes more about using SGP4 ML-corrected predictions - "Improved Orbital Propagator Integrated with SGP4 and Machine Learning"

				mins/rest of class)		
25	26 ASP CLASS Introduction to the Methodology Section rubrics - lecture (45 mins) Update calendar with all due dates for methodology section paper, oral presentation, etc. + format calendar with sectioned work dates (45 mins/rest of time)	27 Continued: format calendar + section work dates for different days	28 ASP CLASS Meeting w Dr. Garcia (10 mins) Started working on changing my dataset to tailor to 1 TLE propagated for a 2 week epoch to feed into my ML model (75 mins/rest of class)	29	30 ASP CLASS **LOGBOOK DUE TODAY** Continued: format calendar + section work dates for different days (30 mins) Cont. work on changing my dataset to tailor to 1 TLE propagated for a 2 week epoch to feed into my ML model (60 mins)	31

Monthly Goals - January

- Logbook due on Jan. 30
  - Plan out February logbook (calendar)
- Finish building input data set in Python to feed into the ML model -> combine Kp, F10.7, SGP4 propagated states, TLE data, epoch, and residuals
- Continue the machine learning process in TensorFlow -> propagate 1 TLE over a period of time and train the ML model to observe the error growth patterns and fix this
- Plan Feb. meeting times

🗓️ February 2026

FEBRUARY 2026

SUN	MON	TUE	WED	THU	FRI	SAT
1	2 Send priority email to mentor (SGP4 manipulation or just residual correction)	3 ASP CLASS Cont. work on changing my dataset to tailor to 1 TLE propagated for a 2 week epoch to feed into my ML model (60 mins) Build methodology paper skeleton (create headings and consolidate all sections needed) (30 mins)	4 Make mentor meeting agenda	5 ASP CLASS Meeting w mentor at 9:00am Send meeting summary + next steps email (10 mins) Meeting w Dr. Garcia (20 mins) Work on Methodology paper intro writing (overview + materials/tools inventory in a paragraph format) (15 mins) Start official Methodology paper intro writing (30 mins/rest of time)	6 Basic CYSF portal information due	7
8	9 ASP CLASS Go over Intro section paper advice + comments for general writing/references comments (30 mins) Continue writing methodology section paper (citing data sources + how I converted them into tables) (60 mins)	10 Make mentor meeting agenda Continue writing methodology section paper ("truth error"/residual error formula and explanation)	11 ASP CLASS Meeting w mentor at 1:00pm Send meeting summary + next steps email (10 mins) Continue writing methodology section paper (ML model process + exact indicators) (45 mins) Create background/research question section for poster	12 **METHODOLOGY SECTION DRAFT DUE TODAY** Create methodology section for poster	13 LONG WEEKEND	14 Finish results/analysis sections for poster

15	16	17	18	19	20	21
Finish results/analysis sections for poster	**ORAL PRESENTATION/ POSTER #2 DRAFT DUE TODAY** LONG WEEKEND	ASP CLASS Incorporate methodology section paper feedback into paper + references check (45 mins) Update CYSF portal with up-to-date information (30 mins) Finalize poster design so I can print over the weekend (15 mins) LUNCH (11:45am): Oral presentation practice today	Make mentor meeting agenda	ASP CLASS **METHODOLOGY SECTION PAPER DUE TODAY -> MOVED TO MARCH 9TH!** Meeting w mentor at 10:00am Send meeting summary + next steps email (10 mins) Add more info to CYSF portal (15 mins) Finalize poster design and incorporate OP #2 edits into poster (30 mins/rest of time)	Practice for OP #2	Practice for OP #2 + PRINT POSTER
22	**23**	**24**	**25**	**26**	**27**	**28**
Practice for OP #2 + PRINT POSTER	**ORAL PRESENTATION #2 CLASS 1 (MY PRESENTATION DAY)** ASP CLASS Listen to 3x presentations + present! (90 mins/whole class)		**ORAL PRESENTATION #2 CLASS 2** ASP CLASS Listen to 1x presentations (20 mins) Send mentor an email to plan MARCH MEETING TIMES (10 mins) Finalize February logbook + add/plan March calendar information (45 mins/rest of class)		**ORAL PRESENTATION #2 CLASS 2** **LOGBOOK DUE TODAY** ASP CLASS Finalize February logbook + add/plan March calendar information (30 mins) Add in my Results Data Collection/Test Results and Discussion/Conclusions data into my logbook tabs (60 mins/rest of	Practice for school Science Fair on March 2

					class)	

Monthly Goals - February

- Logbook due on Feb. 27
    - Plan out March logbook (calendar)
- Finish ML process and gather results for data analysis
- Finish data analysis of ML model in Python
- Finish Methodology Section paper -> due on Feb. 19 -> CHANGED TO Mar. 9th
- Finish Oral Presentation #2 (Science Fair practice) -> on Feb. 23
- Finish Poster outline -> around Feb. 27
- Prep for school Science Fair on March 2
- Plan March meeting times

🗓️ March 2026

MARCH 2026

SUN	MON	TUE	WED	THU	FRI	SAT
1 Prepare + practice presentation for school Science Fair tomorrow	2 SCIENCE FAIR	3 SCIENCE FAIR RESULTS **CYSF PORTAL DUE TONIGHT** ASP CLASS Finish CYSF portal (all information) (60 mins) Incorporate Dr. Garcia edits into my methodology section paper (30 mins/rest of class)	4 Make mentor meeting agenda Continue working on: ML validation and Statistical Analysis of Results portions of my methodology section paper	5 ASP CLASS Meeting w mentor at 9:00am Send meeting summary + next steps email (10 mins) Finish working on ML validation and Statistical Analysis of Results portions of my methodology section paper (60 mins/rest of class) End of Day: Send Dr. Garcia + my mentor my final Methodology Section paper draft for edits/advice	6 Continue editing methodology paper + incorporate Dr. Garcia/mentor edits	7 Continue editing methodology paper + incorporate Dr. Garcia/mentor edits
8 Continue editing methodology paper + incorporate Dr. Garcia/mentor edits	9 **METHODOLOGY SECTION PAPER DUE TODAY** ASP CLASS Tentative - Start future work parts of project: Make a new ML dataset using a longer propagation period (28 days) and more TLE packets (90 mins/whole class)	10 Make mentor meeting agenda	11 ASP CLASS Meeting w mentor at 12:30pm Send meeting summary + next steps email (10 mins) Tentative - Finish new ML dataset and making a new Pandas Dataframe (for the CASSIOPE satellite dataset v2) (30 mins) Tentative -	12	13 ASP CLASS Tentative - Calculate residual error using the along track/across track/radial coordinate frame (60 mins) Tentative -Start drafting my Results paper (skeleton structure)? (30 mins)	14

			Research into new coordinate system (along-track, across-track, radial) + take notes on the benefits/drawbacks of using it (60 mins/rest of class)			
15	16	17 ASP CLASS Tentative - Start ML process for CASSIOPE dataset v2 using the same 80%/20% ML training/validation split that I used for CASSIOPE dataset v1 (60 mins) Finalize March logbook + add/plan April calendar information (30 mins/rest of class)	18 Make mentor meeting agenda Plan April meeting times	19 **LOGBOOK DUE** ASP CLASS Meeting w mentor at 12:30pm Send meeting summary + next steps email (10 mins) Finalize March logbook + add/plan April calendar information (45 mins) Tentative - Continue fleshing out skeleton sections of results section paper? (45 mins/rest of class)	20	21 SPRING BREAK
22 SPRING BREAK	23 SPRING BREAK	24 SPRING BREAK	25 SPRING BREAK	26 SPRING BREAK	27 SPRING BREAK	28 SPRING BREAK

29	30	31	
SPRING BREAK	SPRING BREAK	SPRING BREAK	

Monthly Goals - March

- Logbook due on March 19th
  - Plan out April logbook (calendar)
- School Science Fair on March 2nd -> practice oral presentation
- Finish CYSF portal (all information) -> due on Mar. 3rd
- Finish Methodology Section paper -> due on Mar. 9th
- Tentative - Start "future work" tasks for my project
  - Making a new dataset with more extensive information about CASSIOPE, and increasing the propagation time from 14 days to 28 days
  - Calculating residual error using the along track/across track/radial coordinate frame instead of the current x/y/z Cartesian plane coordinate frame
- Tentative - Writing my results section paper -> due ?
- Plan April meeting times

🗓️ April 2026

APRIL 2026

SUN	MON	TUE	WED	THU	FRI	SAT
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30		

Monthly Goals - April

- Logbook due on April BLANK

🗓️ May 2026

MAY 2026

SUN	MON	TUE	WED	THU	FRI	SAT

Monthly Goals - April

- Logbook due on April BLANK

🌅 Summer 2025 Work

SUMMER 2025 WORK

Personal Research:

- I did lots of research online into machine learning and how AIs "teach themselves" or reinforce their learning through various methods
- I subscribed to many YouTube channels so I could keep up with recent changes in the AI field
- Listened to AI-related podcasts and StarTalk by Neil deGrasse Tyson for extra information
- Summer research and rough notes document: Summer Research & Rough Notes
- I am currently taking lessons on AI and Machine Learning on Brilliant.org

Coding Skills:

- Over the summer, I used ChatGPT 5 Thinking (with my "Plus" subscription) to generate a learning Python course *specifically tailored to my project requirements*
    - Some topics of study include: Python coding basics, Python libraries, organizing data sets using Python, machine learning, AI neural networks
    - Course syllabus: ChatGPT 5 Python Course Syllabus
    - I am currently on unit 3/8

Mentor Meeting:

- Introduction to mentor (Mr. Andrew Howarth) through a Zoom meeting
- Discussion about his involvement in the University of Calgary's **CASSIOPE/e-POP Mission**

    - *From Andrew Howarth's UCalgary Profile*: Launched in 2013, the e-POP payload on the CASSIOPE satellite has been collecting data from low-earth orbit to help scientists study the impact of solar storms on space weather, including auroras, low-energy particle trajectories, radio wave propagation, GPS signals, satellite drag, and more. The data has also been used to locate and track space debris using both optical sensors and plasma wave sensors. The data are openly available from https://epop.phys.ucalgary.ca.
- Discussed alternative project idea (auroral classification using an AI system)
- Mr. Howarth sent me some resources for future reference + as a starting point for research at the beginning of the school year
    - Fast Auroral Imager (FAI) Article
    - CASSIOPE Article
    - FAI Extra Info
    - The Essential Main Ideas of Neural Networks

📓 Daily Notes

DAILY NOTES

See individual months.
For SPECIFIC INFORMATION about tasks completed, time taken to complete them, and upcoming due dates, check the Calendar tab.

August/September Daily Notes

DAILY NOTES - AUGUST/SEPTEMBER

Sep 29, 2025

Class #11: {Logbook Due Today} Work Block

- Went over some PaperPile class questions
  - Sometimes have to copy the DOI separately
  - Can usually add research paper from a google search result or by clicking into the website/journal
- Reminder: 3 main logbook parts
  1) Content (daily notes and background research)
  2) Structure (organization of different tabs)
  3) Calendar (plan out this current month and the next month)
- Oral presentation on Oct. 22/24/28
  - Content grade (out of 34)
  - Participation grade (out of 10)

Sep 25, 2025

Class #10: Work Block

- Main focus for today: FINDING A SPECIFIC RESEARCH QUESTION THAT I CAN PRESENT TO MY MENTOR SO I CAN START WRITING MY RESEARCH PROPOSAL
- Make a mentor meeting agenda (see September Mentor Meeting notes tab)
- Import the new article my mentor retrieved for me ("A Machine Learning-Based Approach for Improved Orbit Predictions of LEO Space Debris With Sparse Tracking Data From a Single Station") to paperpile

    - Start summary + annotations on (see September Background Research tab)
- Research into LeoLabs briefly to understand the commercial applications of the AI and orbital debris field
- Read and make notes on ML-based Approach to Improving Orbital Predictions research paper (see September Background Research tab)

Sep 23, 2025

Class #9: Work Block

- Paperpile does not support documents with tabs
    - I need to make a separate document for my R.P.
- Dr. Garcia says that I should start drafting my R.P. -> I can show it to her to get advice + further improvements + identify gaps in my knowledge
- Starting reading and making notes on ML-Approach to Classifying Orbital Debris research paper (see September Background Research tab)
- Falling a bit behind on timing -> reading and taking notes on research papers takes longer than expected
    - Only had time to start a brief outline on how I am doing to draft my R.P.

Sep 19, 2025

Class #8: Work Block & Dr. Garcia Meeting

- MAKE SURE TO STAY ON TOP OF EMAILS!
- Respond to Mr. Howarth's emails about future meeting times and new research papers
- Dr. Garcia Meeting:
    - Discussed progress done since last meeting with her
    - Discussed how I got more clarification on direction of project + current tasks (deciding what data source I will use - image/TLE/CVS data)
    - Talked about making a FLOW CHART to confirm methodology + give to my mentor
    - Talked about next steps in my project -> starting to rough draft my R.P. w/ intro and methodology
- Took notes on the ESA Annual Report (see September Background Research tab)
- Started taking notes on RSO detection paper (see September Background Research tab) + searching up unknown terms

Sep 17, 2025

Class #7: Work Block & Mentor Meeting

- Meeting with mentor at 10:00 am (see September Mentor Meeting Notes tab)
- Sent meeting follow-up email
- Scanned various research papers (listed in September Background Research tab) for extension questions to find the niche question my project will try to answer
    - Used Google Scholar as a resource to find new research papers
    - Added all the papers into my PaperPile library
    - Mainly scanned the "Further Extensions" portions of research papers to get extension ideas that I can use for my project idea

Sep 15, 2025

Class #6: Work Block & Dr. Garcia Meeting

- Tips given in class today:
    - Google terms and basic info
    - Look for free resources - gather paid sources to bring to mentor
    - Find lots of PEER-REVIEWED papers (google scholar, research gates, pubmed)
    - R.P. Outline:
        - Broader topic -> specific topics -> specific research question
- Dr. Garcia Meeting:
    - Discussed current state of the project + I need to do lots of background research into machine learning algorithms so I get a more in-depth understanding -> can write my research proposal
- Starting watching the "Neural Networks" video and taking notes
- Read over the DASP Meeting presentation + research terms that were unfamiliar to get a better understanding of the subject area

Sep 11, 2025

Class #5: Mentor Meeting & PaperPile Set-Up

- Meeting with mentor today at 12:30 pm (see September Mentor Meeting tab)
- Set-up PaperPile:
  - Pick the citation style that is in most of research papers I am using
  - Make sure to check citations with Owl Purdue to check
  - IEEE citation style for astrophysics
  - Reference list at the bottom of the R.P.
  - In-text citations go on CHRONOLOGICAL ORDER
  - Chrome extension -> can directly share a page from the web
  - Digital annotation with comments and highlights
  - Can export background research + annotations into logbook
  - Put mainly published PEER REVIEWED SOURCES into PaperPile
  - On Google Docs…
    - PaperPile option is on the top menu
    - When insert citation, can select a source from the library/folders
    - Engineering citation = IEEE
    - PAPERPILE DOES NOT WORK WELL WITH DOCS TABS
- OwlPurdue for research and citation specifics
  - https://owl.purdue.edu/owl/
  - Will use APA 7th Edition or IEEE style -> need to ask my mentor for clarification

Sep 9, 2025

Class #4: Work Block

- Reminders:
  - Installing PaperPile next class
  - Make sure to have lots of background research for the rough draft of the R.P.
  - Research questions/goals/methods all included in R.P.

- Make a detailed outline of the calendar I will be using for my logbook and created a Calendar tab with it
- Brainstormed some ideas of things I will talk with my mentor about at my meeting on Thursday (next class)
    - Check September Mentor Meeting Notes tab for specific questions

Sep 5, 2025

Class #3: Work Block

- Ideas of things to work on during work blocks
    - Logbooks:
        - Edit mentor meeting notes
        - Edit calendar
    - Email:
        - Mentors
        - Teachers
    - Research
        - Reading papers
            - Can use generic google searches to familiarize then use published, reputable papers
            - Engineering website source: **IEEE**
            - Next class, we will download PaperPile (web-based Reference Manager) to annotate research papers
        - Research Proposal outline
            - Establish intro + research question
- If research papers are NOT PMC-free downloads, ask mentor for UCID (University of Calgary ID card)
- **Important: a week before deadline, send paper/project to mentor before submitting**

Sep 3, 2025

Class #2: Logbook Guidelines & Intro to Research Proposal

- **Discussion about logbooks**
    - If using Google Docs…
        - Use tabs, colour code tasks, be specific
        - USE DATES FOR EVERYTHING
    - If using Notion…
        - Good for organization but might have to pay for putting a lot of information in
    - Calendar
        - Specifically describe what I will do in every class
    - Research
        - State what pages were read, date read, title and author of the article
        - “Paper Pile” software will be introduced next week to annotate articles
    - September logbook check will be posted on GC today
    - Be proactive to present logbook design to Dr. Garcia and ask what she wants me to add
- **Grading**
    - Bi-weekly check grades from email communications w teachers + mentor, preparedness in all meetings, individual checks
    - Monthly mentor marks
    - **Science Fair:**
        - Gold medal - 100%
        - Silver medal - 95%
        - Bronze medal - 90%
        - HM - 80%
- MARKS LOCK ON THE WEEK OF NOVEMBER 14TH
- **Discussion about research proposal (first major assignment)**
    - If an innovation project -> need to show a CLEAR PROBLEM and propose a solution
    - Make sure research proposals are not TOO broad and also doesn’t take too long
    - Study something that is important in the real-world
    - Good research proposal = funding, interest from scientists, beginning research
    - Guidelines for research proposal on GC
        - **Main points:** TITLE CANNOT BE TOO NICHE, spending a lot of time on the introduction will benefit later, can skip variables in

innovation project, goals = immediate short-term goals (can be achievable this year) & long-term goals (over the course of many years), how does the proposed method help in the investigation?

Aug 29, 2025

Class #1: Class Expectations Lecture

Notes:

- **My teacher mentor: Dr. Garcia**
- Planning
- Time management and organization
    - NO PROCRASTINATION
    - Always add dates to research and daily notes
- Track your task
- Communication w mentors and teachers
    - **Weekly summary emails for teachers?**
    - **Weekly sessions with mentor**
    - CC Ms. Kale and Dr. Garcia while communicating with mentor
    - Email attend@webberacademy.ca when going off campus during ASP (parents email) CC teachers
- Reply as soon as possible -> 3 cheeks per day
- **Personal email with mentor and school account for teachers**
- Send a summary email after all meetings w mentor (CC teachers)
    - Things discussed, tasks given for a due date, next meeting date
    - Type out notes in logbook in every meeting
    - **Have I missed anything?** at the end of the email
- Do not meet mentors during other subject times
    - VERY FEW EXCEPTIONS
- Typical class:
    - Meetings with mentor in room 303 or off campus
    - Work time in room 301
    - Meetings with Dr. Garcia or Ms. Kale
        - Students listed on Google Calendar
        - Be prepared and ready for meetings w teachers
    - Can do work from other classes *if you are caught up on ASP work*
    - Daily work notes in the logbook

- Assessed:
    - Logbook
        - Data collection
        - Notes from meetings
        - Research
        - Daily notes from class
        - Calendar (for time management -> think about this before next class)
            - Plan for the current month and the next month (tasks for each day in the month + due dates)
            - Plan things **SPECIFICALLY** with timestamps to fill up the 90 minute class
    - Mentor assessment
        - Attendance at meetings
        - Punctuality
        - Meeting deadlines
        - Preparedness
        - Communication
- Evaluation of Class:
    1) Schedule, communication, organization (20%)
        a) Out of 15, biweekly mark
    2) Logbook (20%)
        a) 2% overall for each month, monthly mark at the END OF THE MONTH
    3) Oral Presentations (30%)
        a) 4 major in-class oral presentations
    4) Written Work (30%)
        a) Writing scientific papers and citations, 6 papers

🗓️ October Daily Notes

DAILY NOTES - OCTOBER

Oct 30, 2025

Class #22: {Oral Presentations Day 2, Logbook due today} Mentor Meeting & Watching Presentations

- Meeting with mentor at 10:00 am (see October Mentor Meeting Notes tab)
- Sent meeting follow-up email
- WATCHED PEERS' ORAL PRESENTATIONS TODAY (see Major Assessments -> Oral Presentation tab) -> notes/questions about other presentation
- General advice from Dr. Garcia:
  - Make November calendar and be SPECIFIC
  - Should include <u>SAMPLE SIZE</u> in my oral presentation
  - November task: enter basic project info into the CYSF portal -> make sure to do the ethics consideration form
    - I can declare that I am NOT using humans or animals in my project

Oct 28, 2025

Class #21: {Oral Presentations Day 1} Work Block

- Main working focus today: finish working on 3/3 Intro slides and finish 1/1 research question slides (see Major Assessments -> Oral Presentation tab)
  - Starting planning out the skeleton structure of the rest of my slides for my mentor to approve before my oral presentation on Monday

- General reminder: Logbooks must have notes on the presentations

Oct 24, 2025

Class #20: Work Block & Dr. Garcia Meeting

- Main working focus today: start building OP outline with slide breakdown, find presentation template, and start working on 1/3 of the Intro slides (see Major Assessments -> Oral Presentation tab)
- General tips for the oral presentation
    - Little text and big fonts
    - More GRAPHICS (images, graphs, tables, flow diagrams) -> high quality graph that is clear
    - Should use up the space of the whole slide
    - Citations:
        - Typically do not have to cite information on slides -> do not need [1] by info on the slides
        - Give credit to the source ("work/source/graphic is taken from…")
        - If I make a source myself, I do not need to give credits
        - If a source from a website or mentor, say "courtesy of…"
    - Methodology:
        - Flow diagram is STRONGLY SUGGESTED
        - Give overview of methodology, can do extra slides for some parts to give more information
    - Should start with a slides outline first to plan out which content should go where

Oct 22, 2025

Class #19: {Research Proposal Due Today} Work Block & Mentor Meeting

- Main working focus today: edit R.P. using mentor edits to make sure it is ready to hand in for 11:59pm tonight (see Major Assessments -> Research Proposal tab)
- Meeting with mentor at 10:00 am (see October Mentor Meeting Notes tab)
- Sent meeting follow-up email
- Vague format for the rest of the year:

    - RP (due today)
    - Introduction (Nov-Jan)
    - Methods (Feb)
    - Results (March)
    - Analysis (April)
- School science fair at the beginning of March (poster and presentation)
- Class discussion about oral presentation (+ tips):
    - Tip: start with a punchline/brief summary of project to keep the audience hooked in a way they will understand (like elevator pitch)
    1) Create a presentation outline with 8-10 slides, breaking up RP into sections easy to understand
    2) Breakdown how many slides will be done per day or per class
    3) Practice the presentation to make sure it is within the time limit
        a) Focus on UNDERSTANDING
        b) Focus on TIMING (10 minutes) -> need to rehearse

Oct 20, 2025

Class #18: Work Block

- Main working focus today:
- Finish R.P. significance + references sections, as well as editing the methodology (with flowchart) section of my R.P. (see Major Assessments -> Research Proposal tab)
- Reading and making notes on "Precise and Efficient Orbit Prediction" (see Background Research -> October Background Research tab)
- Need to have a title page
    - Need to include my name, mentor's name and department, research proposal, applied science project, date
- After I finish the research proposal, I need to submit it to Turnitin.com and Google Classroom

Oct 16, 2025

Class #17: Work Block & Mentor Meeting & Dr. Garcia Meeting

- Main working focus today: finding research sources that explains how to use pandas in Python (and using Python DataFrames) and starting work on the methodology section of my R.P. (see Major Assessments -> Research Proposal tab)

- Meeting with mentor at 1:00 pm (see October Mentor Meeting Notes tab)
- Sent meeting follow-up email
- Should use Purdue Owl for specifics for citations and inserting diagrams/figures (IEEE)
- I should check the new guidelines/rubric for the research proposal to see that I have done each section + in the correct order
    - I have an innovation project -> the methodology is out of 30 for me since I do not have to do a hypothesis/variables

Oct 14, 2025

Class #16: Work Block

- Main working focus today: include R.P. edits from Dr. Garcia + finish the objectives section of the R.P. (see Major Assessments -> Research Proposal tab)
- Received feedback from Dr. Garcia
    - Started implementing feedback on intro + research question
- Oral presentation tips
    - Have 10 minutes -> be concise on slides with not too many words on slides
        - Use images and graphics and LESS TEXT (with big font)
        - When taking a figure, then cite it/give credits
        - Information does not need to be cited explicitly
        - Should use a **flowchart for methodology**
        - Pictures can be found in **BIORENDER** for images for methodology or slideshow
            - To cite an image, write “taken from…, kindly provided by…”
    - Use technical terminology, but **define the terms for the audience**
    - Look at rubric for expectations
    - Confirm slideshow with Dr. Garcia before the presentation for feedback
        - Slideshow should not be very flashy and have limited distractions
    - Suggested order of slides
        - **Title (1 slide) w/ “punchline”**
        - **Intro (3/4 slides)**
        - **RQ/Goals/Objectives (2/3 slides)**
        - **Methodology (3/4 slide)**

        - **Significance/Conclusion (1 slide)**
        - **Conclusion/Ending slide (1 slide)**

Oct 9, 2025

Class #15: Work Block & Dr. Garcia Meeting

- Main working focus today:
    - Work in research proposal research question and objectives (see Major Assessments -> Research Proposal tab)
    - Reading and making notes on “Colliding Satellites: Consequences and Implications” (see Background Research -> October Background Research tab)
- General reminders:
    - I should be constantly emailing my mentor (especially with meeting summaries) to figure out what tasks I need to complete and checking if I have all the information correct
- Dr. Garcia meeting notes:
    - Extension on research proposal to Oct. 22, 2025
    - Extension on oral presentation to Oct 28, 2025
    - Should take out the variables section of my research proposal -> innovation project (machine learning variables are too complex to include right now)
    - Should restate the problem + need for solving problem with **research question at the END OF THE INTRODUCTION**
    - Research question/hypothesis section after introduction
    - Objectives should have **short-term and long-term goals**

Oct 7, 2025

Class #14: Work Block

- Meeting with Ms. Parker from 12:50pm to 1:20pm (lost some work time)
- Main working focus today: worked on the finishing the first draft of my research paper introduction (see Major Assessments -> Research Proposal tab)
- Must join CYSF online portal for science fair
    - Put in basic project information and ethics proposal (if necessary)
- Work time on research proposal introduction and objectives

Oct 3, 2025

Class #13: Work Block & Mentor Meeting

- Meeting with mentor at 10:00 am (see October Mentor Meeting Notes tab)
- Sent meeting follow-up email
- Main working focus for today: work on R.P. intro + research questions (see Major Assessments -> Research Proposal tab)
- Dr. Garcia general lecture to the class:
  - Mentors will give a monthly evaluation (3 sections -> progress, communication, readiness) out of 15 total marks
    - Reflect on all aspects/areas of improvement
  - Remember to plan tasks for future months
  - For R.P.:
    - If you need to cite info from the same source for a whole paragraph, you must cite [1] at the beginning and [1] at the end of the paragraph
    - **Must cite the primary source and secondary source (if applicable) -> MUST ADD MULTIPLE PAPERS TO THE CITATION**
    - If using a figure, must use the style of IEEE (**figure on top**, “Fig 1.”, title, description)
      - Should not have a figure in one page with a figure legend in the next page (keep them together!)

Oct 1, 2025

Class #12: Work Block

- Main working focuses for today:
  - Work on Intro/Title for R.P. (see Major Assessments -> Research Proposal tab)
  - Reading and making notes on “Improving Orbit Prediction Accuracy through Supervised Machine Learning” (see Background Research -> October Background Research tab)
- Some general/important feedback of my logbook:
  - Should replace the title of my logbook with the title of my project when I have it

    - Should remember to always specifically mention where to find information I mention in my daily notes -> direct to specific tabs (like Mentor Meeting Notes, Background Research)
    - Good format to gather information about my background research sources -> can maybe add the DOI information?
    - Excellent organization of my mentor meeting notes
    - Good colour organization, specific tasks, and monthly goals in my calendar
- General reminders from in class today:
    - Should cross out/colour code tasks that have been completed -> I am already doing this
    - I can submit the link to my logbook so I can keep working on it while it is being graded by Dr. Garcia
    - Should have a to-do checklist with due dates-> I have this in my calendar/mentor meeting notes tabs

🗓️ November Daily Notes

DAILY NOTES - NOVEMBER

Nov 27, 2025

Class #31: MISS CLASS BECAUSE OF VBALL PROVINCIALS {LOGBOOK DUE TODAY}

Nov 25, 2025

Class #30: MISS CLASS BECAUSE OF PRESIDENT'S BREAKFAST

Nov 21, 2025

Class #29: Work Block

- Main points from Dr. Garcia lecture today:
    - We should be constantly pushing towards our deadlines
    - Always keep communication with mentors (even if we are away or they are away)
        - Dr. Garcia should see emails from us every week in her inbox
        - Remember to always CC Dr. Garcia in our emails
- Main working focus for today:
    1) Start inputting our downloaded project data into Python DataFrame
        a) Using the Pandas library and the import pandas as pd function
        b) Used similar code to the Cisco online course + used this to guide me through the importation steps

2) Background research about how to use SPG4 integrated into Python - CONT.
    a) See (see Background Research -> November Background Research tab) for background research notes about this

Nov 19, 2025

Class #28: Mentor Meeting & Work Block

- Meeting with mentor at 10:00 am (see November Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
    1) Read articles about using Pandas in Python + take notes CONT.
        a) See (see Background Research -> November Background Research tab) for background research notes about this
    2) Practice implementing sample data sets into Python to learn how to create DataFrames - data from Cisco online course (60 mins)

Nov 17, 2025

Class #27: Dr. Garcia Meeting & Work Block

- Main working focus for today:
    1) Research methods to propagate orbits using an SGP4 system in Python
        a) Learned that there are many ways to implement an SGP4/types of systems in Python that can be used to propagate the future orbits of satellites
        b) See (see Background Research -> November Background Research tab) for background research notes about this
    2) General research on using Pandas in Python for data analysis
        a) See (see Background Research -> November Background Research tab) for background research notes about this

Nov 13, 2025

Class #26: Mentor Meeting & Work Block

- Meeting with mentor at 12:30 pm (see November Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
    1) Finish data acquisition of all types of data + space weather data (line these up by Jan 01, 2019 - June 30, 2019)
    - Finished downloading TLE data from space-track.org
    - Finished downloading GPS (SP3s) data from https://edex.phys.ucalgary.ca/
    - Finished downloading F10.7cm data from https://lasp.colorado.edu/lisird/data/penticton_radio_flux
    - Finished downloading Kp data from https://kp.gfz.de/en/data
- General lecture notes from Dr. Garcia
    - Only have around 10 weeks left to complete the project -> **need to make a LONG RANGE plan with my mentor at my next meeting with him with specific tasks**
        - Nov - 2 weeks, Dec - 3 weeks, Jan - 2 weeks, Feb - 2 weeks (last 2 weeks for poster + presentation)
    - May need to create an "exit plan" and limit the ambition of the project so it can logistically be done in the time frame available
    - On Winter Break we can make a choice whether we want to work on the project or not -> do not need to meet with mentor but can talk to him about that

Nov 7, 2025

Class #25: {Oral Presentations Day 4} Work Block

- Watched Ronald's presentation today for curiosity (did not take notes)
- Main working focus for today:
    1) Data acquisition of GPS truth data at the epochs of study/truth data ephemeris (60 mins) for CASSIOPE
    2) Started research into space weather data types that I will use for the project -> Kp, F10.7 (30 mins) - notes - (see Background Research ->

November Background Research tab)

- General lecture notes from Dr. Garcia
    - Should be having mentor meeting every week, and send a summary every week with work done + next steps
    - Calendar MUST have 90 minute tasks listed
    - Most data collection and running tests should be done in November (December-January will not have that much work being done because of holidays and Midterms)
    - Should discuss a long-term timetimes, long-term objectives for my project
    - Next paper will be in Mid-December, November is allocated for data collection and analysis

Nov 5, 2025

Class #24: Mentor Meeting & Work Block

- Class lecture today general points:
    - Should be having weekly emails to my mentor -> increases biweekly communication mark
        - Include a progress report
        - Summary notes of the meeting
        - Agenda for next meeting (optional)
        - Plan for next week
    - Need to continue taking notes on class lectures so that logbook is comprehensive
    - Should input the exact section of a project/written assignment in my daily notes, highlighting the work that I did THAT CLASS
- Should be spending 10-15 minutes
- Meeting with mentor at 9:00 am (see November Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
    1) CYSF portal ethics consideration document submitted + finished fully setting up everything about my project in the portal
    2) Started data acquisition of TLEs by creating an account for space-track.org and finding a time period + downloading the TLEs for 6 months in 2019
        - I had trouble with formatting the TLE downloads at a txt file -> research online how to do this

Nov 3, 2025

Class #23: {Oral Presentations Day 3} & Watching Presentations

- Completed my oral presentation in front of the class
- Main focus for today: watched my peers' oral presentations + took notes on them + asked questions (see Major Assessments -> Oral Presentation #1 tab)
    - Watched oral presentation #3 (Shaayaan) and #4 (Samir)

🗓️ December Daily Notes

DAILY NOTES - DECEMBER

Dec 19, 2025

Class #39: Work Block & Dr. Garcia Meeting

- Main working focus for today:
  - Finalize logbook for December + make January calendar (see December/January Calendar tabs)
  - Plan generic February calendar tasks to plan out the tentative rest of the project timeline (TensorFlow portion + analysis) -> reference my email with my mentor where I planned the schedule for the rest of my project this year (see February Calendar tab)
  - Study for physics unit test (remaining time) today

Dec 17, 2025

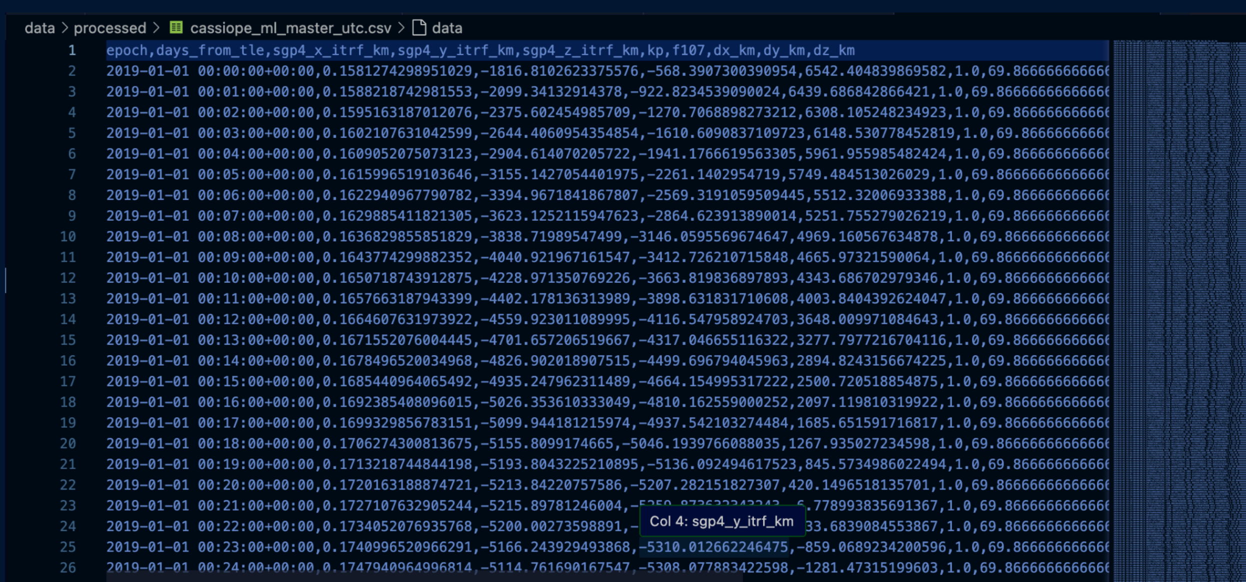
Class #38: Mentor Meeting & Work Block

- Meeting with mentor at 12:30 pm (see December Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
  - Finalize logbook for December + make January calendar (see December/January Calendar tabs)
    - Filled in all the missing details from December daily notes
    - Properly formatted my December background research (see December Background Research tab)
    - Planned January calendar and briefly Winter Break tasks -> finish propagations and learn more about how to integrate

Dec 15, 2025

Class #37: {Introduction Section Paper due today} Work Block & Dr. Garcia Meeting

- For my mentor meeting on Wednesday:
  - I need to clearly outline what work I will be doing over the Winter Break
  - Plan out December and January clearly
  - Meeting plan over the Break and in January after midterms?
- General notes from Dr. Garcia lecture today:
  - Logbook check will be in January for December/January logbook -> need to plan January this week or over Winter Break
- Main working focus for today:
  - Start building the specific Ml-ready dataset and save in Python (see the image below)
    - Save different types of data in the ephemeris (such as epoch, sgp4 propagations in newly converting ITRF datetime, and residual errors)


data > processed > cassiope_ml_master_utc.csv > data

```
epoch,days_from_tle,sgp4_x_itrf_km,sgp4_y_itrf_km,sgp4_z_itrf_km,kp,f107,dx_km,dy_km,dz_km
2019-01-01 00:00:00+00:00,0.1581274298951029,-1816.8102623375576,-568.3907300390954,6542.404839869582,1.0,69.86666666666
2019-01-01 00:01:00+00:00,0.1588218742981553,-2099.34132914378,-922.8234539090024,6439.686842866421,1.0,69.86666666666666
2019-01-01 00:02:00+00:00,0.1595163187012076,-2375.602454985709,-1270.7068898273212,6308.105248234923,1.0,69.86666666666
2019-01-01 00:03:00+00:00,0.1602107631042599,-2644.4060954354854,-1610.6090837109723,6148.530778452819,1.0,69.86666666666
2019-01-01 00:04:00+00:00,0.1609052075073123,-2904.614070205722,-1941.1766619563305,5961.955985482424,1.0,69.86666666666
2019-01-01 00:05:00+00:00,0.1615996519103646,-3155.1427054401975,-2261.1402954719,5749.484513026029,1.0,69.86666666666666
2019-01-01 00:06:00+00:00,0.1622940967790782,-3394.9671841867807,-2569.3191059509445,5512.32006933388,1.0,69.86666666666
2019-01-01 00:07:00+00:00,0.1629885411821305,-3623.1252115947623,-2864.623913890014,5251.755279026219,1.0,69.86666666666
2019-01-01 00:08:00+00:00,0.1636829855851829,-3838.71989547499,-3146.0595569674647,4969.160567634878,1.0,69.866666666666
2019-01-01 00:09:00+00:00,0.1643774299882352,-4040.921967161547,-3412.726210715848,4665.97321590064,1.0,69.86666666666666
2019-01-01 00:10:00+00:00,0.1650718743912875,-4228.971350769226,-3663.819836897893,4343.686702979346,1.0,69.866666666666
2019-01-01 00:11:00+00:00,0.1657663187943399,-4402.178136313989,-3898.631831710608,4003.8404392624047,1.0,69.86666666666
2019-01-01 00:12:00+00:00,0.1664607631973922,-4559.923011089995,-4116.547958924703,3648.009971084643,1.0,69.866666666666
2019-01-01 00:13:00+00:00,0.1671552076004445,-4701.657206519667,-4317.046655116322,3277.7977216704116,1.0,69.86666666666
2019-01-01 00:14:00+00:00,0.1678496520034968,-4826.902018907515,-4499.696794045963,2894.8243156674225,1.0,69.86666666666
2019-01-01 00:15:00+00:00,0.1685440964065492,-4935.247962311489,-4664.154995317222,2500.720518854875,1.0,69.866666666666
2019-01-01 00:16:00+00:00,0.1692385408096015,-5026.353610333049,-4810.162559000252,2097.119810319922,1.0,69.866666666666
2019-01-01 00:17:00+00:00,0.1699329856783151,-5099.944181215974,-4937.542103274484,1685.651591716817,1.0,69.866666666666
2019-01-01 00:18:00+00:00,0.1706274300813675,-5155.8099174665,-5046.1939766088035,1267.935027234598,1.0,69.86666666666666
2019-01-01 00:19:00+00:00,0.1713218744844198,-5193.8043225210895,-5136.092494617523,845.5734986022494,1.0,69.86666666666
2019-01-01 00:20:00+00:00,0.1720163188874721,-5213.84220757586,-5207.282151827307,420.1496518135701,1.0,69.86666666666666
2019-01-01 00:21:00+00:00,0.1727107632905244,-5215.89781246004,-5250.8736...6.778993835691367,1.0,69.866666666666
2019-01-01 00:22:00+00:00,0.1734052076935768,-5200.00273598891,-...33.6839084553867,1.0,69.86666666666
2019-01-01 00:23:00+00:00,0.1740996520966291,-5166.243929493868,-5310.012662246475,-859.0689234200596,1.0,69.86666666666
2019-01-01 00:24:00+00:00,0.1747940964996814,-5114.761690167547,-5308.077883422598,-1281.47315199603,1.0,69.866666666666
```

Dec 11, 2025

Class #36: Work Block & Mentor Meeting

- IMPORTANT NOTE: the CYSF portal closes on March 4th, 2026
- Meeting with mentor at 9:00 am (see December Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
    - Fix TLE propagation by converting to ITRF using Python on VS Code
        - I used a Python program and the Skyfield Python library to convert from TEME to ITRF
        - Resulting propagations decreased
    - Extended research into TensorFlow + how to build the input data set - research paper: "Analysis of the Application Efficiency of TensorFlow and PyTorch in Convolutional Neural Network" (see December Background Research tab)
    - More edits to my introduction paper from Dr. Garcia feedback from research proposal

Dec 9, 2025

Class #35: Work Block & Dr. Garcia Meeting

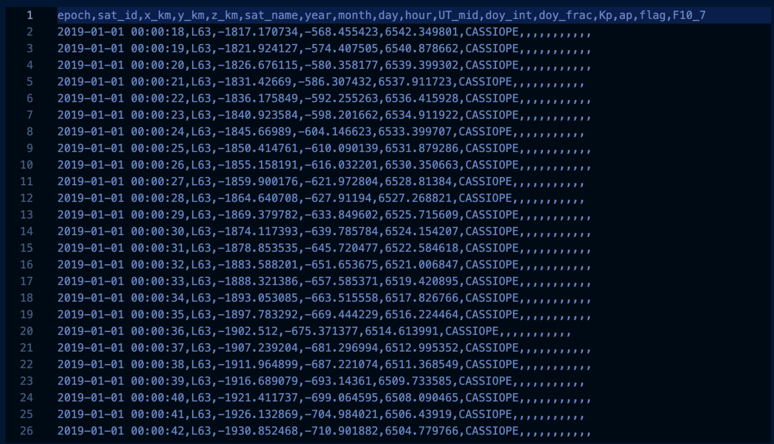
- Notes from general Dr. Garcia lecture today:
    - Continue communication with mentor
    - I should discuss with my mentor if I need to meet with my mentor in Dec/Jan
        - TALK ABOUT PLAN FOR DECEMBER AND JANUARY IN NEXT MENTOR MEETING + CC THIS EMAIL TO DR. GARCIA
    - I will not have individual meetings with Dr. Garcia from **Jan 5th to Jan 20th**
        - Need to continue communication with my mentor though!
        - Need about weekly mentor emails to update them about progress + set up mentor meeting dates
    - If I do not have any concrete tasks to complete, I should continue reading research papers + literature + website about my project topic

- Main working focus for today:
    - Propagate the orbit of 1 TLE and then send this TLE to my mentor
        - Encountered an issue with the TLE propagations since the error was extremely large -> are the propagations in different coordinate systems?
        - I need to discuss this with my mentor in my meeting on Dec. 11

Dec 5, 2025

Class #34: Work Block

- Main working focus for today:
    - Finish working on inputting downloaded data into my Python DataFrame with clear headings and titles ->
    - Make edits on my introduction section paper from Dr. Garcia's feedback on my research proposal
    - Quick research and notes about TensorFlow and PyTorch + their differences - research paper: "TensorFlow: A system for large-scale machine learning" (see December Background Research tab)



```
epoch,sat_id,x_km,y_km,z_km,sat_name,year,month,day,hour,UT_mid,doy_int,doy_frac,Kp,ap,flag,F10_7
2019-01-01 00:00:18,L63,-1817.170734,-568.455423,6542.349801,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:19,L63,-1821.924127,-574.407505,6540.878662,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:20,L63,-1826.676115,-580.358177,6539.399302,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:21,L63,-1831.42669,-586.307432,6537.911723,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:22,L63,-1836.175849,-592.255263,6536.415928,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:23,L63,-1840.923584,-598.201662,6534.911922,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:24,L63,-1845.66989,-604.146623,6533.399707,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:25,L63,-1850.414761,-610.090139,6531.879286,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:26,L63,-1855.158191,-616.032201,6530.350663,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:27,L63,-1859.900176,-621.972804,6528.81384,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:28,L63,-1864.640708,-627.91194,6527.268821,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:29,L63,-1869.379782,-633.849602,6525.715609,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:30,L63,-1874.117393,-639.785784,6524.154207,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:31,L63,-1878.853535,-645.720477,6522.584618,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:32,L63,-1883.588201,-651.653675,6521.006847,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:33,L63,-1888.321386,-657.585371,6519.420895,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:34,L63,-1893.053085,-663.515558,6517.826766,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:35,L63,-1897.783292,-669.444229,6516.224464,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:36,L63,-1902.512,-675.371377,6514.613991,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:37,L63,-1907.239204,-681.296994,6512.995352,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:38,L63,-1911.964899,-687.221074,6511.368549,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:39,L63,-1916.689079,-693.14361,6509.733585,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:40,L63,-1921.411737,-699.064595,6508.090465,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:41,L63,-1926.132869,-704.984021,6506.43919,CASSIOPE,,,,,,,,,,,
2019-01-01 00:00:42,L63,-1930.852468,-710.901882,6504.779766,CASSIOPE,,,,,,,,,,,
```

Dec 3, 2025

Class #33: Mentor Meeting & Dr. Garcia Meeting & Work Block

- Meeting with mentor at 8:30 am (see November Mentor Meeting Notes tab)
- Sent meeting summary email
- Class lecture today general notes:
    - I need to reference the exact Python libraries + databases + tools that I use for the project! -> need to add this into my methodology paper
    - Need to utilize the edits to my research paper introduction from Dr. Garica before I submit my introduction section paper

    - Should keep being specific in my logbook + give SPECIFIC tasks in my logbook
    - Will have introduction paper due on Dec. 15
        - I can use the introduction I used for the research proposal, while also adding in my goals/research questions + ADDING IN DR. GARCIA'S FEEDBACK
        - I should send a draft and send it to Dr. Garcia for her to check + give comments (needs to be done a week before the due date - a week before Dec. 15th)
    - Will have methodology papers due on
        - We will go over how to write this in class later on
- Main working focus for today:
    1) Continue working on inputting downloaded data into Python DataFrame

Dec 1, 2025

Class #32: MISS CLASS BECAUSE OF ASSEMBLY

🗓️ January Daily Notes

DAILY NOTES - JANUARY

Jan 30, 2026

Class #46: Work Block {LOGBOOK DUE TODAY}

- General notes from Dr. Garcia lecture today:
    - Start editing the CYSF portal -> should but this in my calendar for February
    - Need to constantly check my emails for emails from my mentor so I make sure that I reply promptly
    - Reminder: my oral presentation is on Feb. 23rd
    - Other oral presentation dates: Feb. 25th and Feb. 27th
        - I need to attend at least 4 other presentations to get my full participation mark
    - Around 10-12 minute presentations
    - For this presentation, instead of just telling my classmates generally what my project is about -> this time I need to talk about my WHOLE project (whole scope) -> be clear and concise and focus on the big picture
    - Planning:
        - Background: 1-2 mins
        - Methodology: 2 mins
        - Results: 4-5 mins
        - Analysis/Conclusions: 2-3 mins
    - Can show my poster planning to Dr. Garcia online/plan on paper before I print it out -> can get feedback and edits
- Main working focus for today:
    - Continue formatting calendar + section work dates for different days in the month of February (see February Calendar tab)

    - Continue work on changing my dataset to tailor to 1 TLE propagated for a 2 week epoch to feed into my ML model
        - For more information about the specifics about this new dataset's restrictions -> (see January Mentor Meeting Notes tab)

Jan 28, 2026

Science Fair Meeting:

- General notes from the Science Fair Meeting today:
    - Must complete Ethics Form 2A (in CYSF portal) -> due Feb. 6th (my project does not need ethics approval)
    - School Science Fair date -> all day on March 2nd (email teachers to tell them I will be missing classes the whole day)
        - Location: PAC on the stage or in the lobby
        - Judging: 2 periods (randomly will have around 5 judges throughout this period - can bring some light homework.reading to do during the waiting times)
            - 9:00 am - 11:30 am
            - 12:30 pm - 3:30 pm
    - Top 15 JH/SH projects will advance to CYSF competition (April 9-11)
    - CYSF portal closes: March 4 (for CYSF participants advancing)
    - The school will provide a trifold, but I may also use my own trifold or a printed poster -> need to let Dr. Garcia know my preference soon
    - Check theCYSF judge sheet for the OP #2 rubric since we will be using the same one in class

Jan 28, 2026

Class #45: Dr. Garcia Meeting & Work Block

- Main working focus for today
    - Started working on changing my dataset to tailor to 1 TLE propagated for a 2 week epoch to feed into my ML model
        - My mentor said that we will instead be working with sets of 1 data package/TLE propagated over a 2 week epoch so that we can observe increase in residual error growth along with the

respective space weather data (see January Mentor Meeting notes)
        - The ML model will be able to make changes to SGP4 parameters using these new training sets of data

Jan 26, 2026

Class #44: Work Block

- General notes from Methodology Section paper lecture today:
    - One of the best ways to learn how to format the Methodology Section of my paper -> read OTHER methodology sections of papers of related works
    - Experimental Procedure (easy-to-follow steps so that other scientists to repeat the experiment successfully, specific materials and equipment used, deep understanding of the methods and equipment employed)
    - Methodology will be explained to experts in the field -> basic ideas do not need to be overexplained
    - Reference the equipment/resources that I used in my methodology (should look at research papers in related fields to see how this is done)
    - Methodology written in past tense -> it is already completed and it will be written in past tense in the final research paper that I will write at the end of the year
    - Materials/methods should be explained in a “narrative”, not a bulleted list, in depth
    - Cite databases
    - Tables and figures are necessary, but if I add them I have to make sure they add to the paper not subtract from its quality
        - All figures need a figure number AND a figure legend
    - The reference list should include peer-reviewed scientific papers and sources
    - Beginning: brief description about aims/goals/summary of the methods used and then expand from there
    - Writing style: technical, dry, to the point
    - I should create different subtitles for the different sections of the methodology to clearly show the different techniques used
    - Purdue Owl have lots of structural information that I should reference when writing my methodology section paper
    - Cite all software and **VERSIONS**

- Main working focus for today:
  - Introduction to the Methodology Section rubrics - lecture
  - Update calendar with all due dates for methodology section paper, oral presentation, etc. + format calendar with sectioned work dates - (see December/January/February Daily Notes and Calendar tabs)

Jan 22, 2026

Class #43: Mentor Meeting & Work Block

- Meeting with mentor at 10:00 am (see January Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
  - Finish building ML-ready ephemeris for TensorFlow ->>>>>
  - Clean space weather data (Kp and F10.7) by sorting the CSV files for epoch in ITRF and resulting space weather value (BELOW)

```
epoch,tle_index_used,days_from_tle,truth_x_km,truth_y_km,truth_z_km,sgp4_x_itrf_km,sgp4_y_itrf_km,sgp4_z_itrf_km,positior
2019-01-01 00:00:18+00:00,2961,0.15833576302975416,-1817.170734,-568.455423,6542.349801,-1902.155456200402,-675.32005055
2019-01-01 00:01:18+00:00,2961,0.15903020789846778,-2099.689359,-922.843738,6439.641417,-2182.9329418315915,-1027.9427651
2019-01-01 00:02:18+00:00,2961,0.1597246523015201,-2375.940057,-1270.683676,6308.074704,-2457.079376017708,-1373.5787252
2019-01-01 00:03:18+00:00,2961,0.16041909670457244,-2644.735431,-1610.543991,6148.519864,-2723.4214390835286,-1710.81890
2019-01-01 00:04:18+00:00,2961,0.16111354110762477,-2904.937322,-1941.071854,5961.968955,-2980.8367960751857,-2038.33463
2019-01-01 00:05:18+00:00,2961,0.1618079855106771,-3155.462073,-2260.99842,5749.525238,-3228.2590041556073,-2354.8827050
2019-01-01 00:06:18+00:00,2961,0.16250242991372943,-3395.28495,-2569.143275,5512.39205,-3464.6816624956737,-2659.3094359
2019-01-01 00:07:18+00:00,2961,0.16319687431678176,-3623.443717,-2864.417742,5251.861602,-3689.161820216207,-2950.55366
2019-01-01 00:08:18+00:00,2961,0.1638913187198341,-3839.041546,-3145.827159,4969.303835,-3900.8226698846647,-3227.648737
2019-01-01 00:09:18+00:00,2961,0.16458576312288642,-4041.249212,-3412.472157,4666.155435,-4098.855564033698,-3489.723551
2019-01-01 00:10:18+00:00,2961,0.16528020752593875,-4229.306627,-3663.549055,4343.909191,-4282.52140010471,-3736.0027715
2019-01-01 00:11:18+00:00,2961,0.16597465192899108,-4402.523814,-3898.349483,4004.103759,-4451.151425172751,-3965.806246
2019-01-01 00:12:18+00:00,2961,0.1666690963320434,-4560.281311,-4116.259303,3648.313911,-4604.147515872595,-4178.547746
2019-01-01 00:13:18+00:00,2961,0.16736354120075703,-4702.03013,-4316.756926,3278.141332,-4740.982004100761,-4373.730816
2019-01-01 00:14:18+00:00,2961,0.16805798560380936,-4827.291308,-4499.411107,2895.205976,-4861.197030527498,-4550.957778
2019-01-01 00:15:18+00:00,2961,0.1687524300068617,-4935.655072,-4663.878279,2501.138001,-4964.403673650412,-4709.904184
2019-01-01 00:16:18+00:00,2961,0.16944687440991402,-5026.779683,-4809.89949,2097.570327,-5050.280659336005,-4850.338359
2019-01-01 00:17:18+00:00,2961,0.17014131881296635,-5100.390023,-4937.297042,1686.13186,-5118.572891318271,-4972.106626
2019-01-01 00:18:18+00:00,2961,0.17083576321601868,-5156.275978,-5045.970904,1268.441342,-5169.089778612545,-5075.131926
2019-01-01 00:19:18+00:00,2961,0.171530207619071,-5194.290666,-5135.894985,846.10182,-5201.7034183499345,-5159.410039092
2019-01-01 00:20:18+00:00,2961,0.17222465202212334,-5214.348558,-5207.113288,420.695695,-5216.3466760597175,-5225.005708
2019-01-01 00:21:18+00:00,2961,0.17291909642517567,-5216.423521,-5259.736032,-6.219667,-5213.011201015155,-5272.04874279
2019-01-01 00:22:18+00:00,2961,0.17361354082822 8,-5200.546821,-5293.935763,-433.115802,-5191.745409816587,-5300.73011864
2019-01-01 00:23:18+00:00,2961,0.17430798523128033,-5166.805111,-5309.943503,-858.496538,-5152.652466984184,-5311.298132
2019-01-01 00:24:18+00:00,2961,0.17500243009999394,-5115.33842,-5308.044951,-1280.900927,-5095.888287100099,-5304.054630
2019-01-01 00:25:18+00:00,2961,0.17569687450304627,-5046.338189,-5288.576787,-1698.905767,-5021.659579006605,-5279.35133
2019-01-01 00:26:18+00:00,2961,0.1763913189060986,-4960.045345,-5251.923085,-2111.127744,-4930.221948785944,-5237.586323
2019-01-01 00:27:18+00:00,2961,0.17708576330915093,-4856.748457,-5198.511864,-2516.225252,-4821.878089962432,-5179.20056
2019-01-01 00:28:18+00:00,2961,0.17778020771220326,-4736.781964,-5128.811791,-2912.899913,-4696.975979573154,-5104.67476
2019-01-01 00:29:18+00:00,2961,0.1784746521152556,-4600.524495,-5043.329041,-3299.897856,-4555.907316233073,-5014.526179
2019-01-01 00:30:18+00:00,2961,0.17916909651830792,-4448.397284,-4942.604322,-3676.010781,-4399.1058988224195,-4909.3057
2019-01-01 00:31:18+00:00,2961,0.17986354092136025,-4280.86267,-4827.210059,-4040.076848,-4227.046177832331,-4789.595131
2019-01-01 00:32:18+00:00,2961,0.18055798532441258,-4098.422706,-4697.747753,-4390.981424,-4040.2418874061705,-4656.0045
2019-01-01 00:33:18+00:00,2961,0.18125242972746491,-3901.617857,-4554.845488,-4727.657714,-3839.2447717018035,-4509.1698
2019-01-01 00:34:18+00:00,2961,0.18194687413051724,-3691.025789,-4399.155609,-5049.087318,-3624.643402192128,-4349.7504
2019-01-01 00:35:18+00:00,2961,0.18264131899923086,-3467.260244,-4231.352539,-5354.300725,-3397.062080979038,-4178.42727
2019-01-01 00:36:18+00:00,2961,0.1833357634022832,-3230.970005,-4052.130745,-5642.377773,-3157.1598239643304,-3995.90061
2019-01-01 00:37:18+00:00,2961,0.18403020780533552,-2982.837925,-3862.202846,-5912.448109,-2905.6294163739467,-3802.8883
2019-01-01 00:38:18+00:00,2961,0.18472465220838785,-2723.580036,-3662.297844,-6163.691659,-2643.196532081613,-3600.12404
2019-01-01 00:39:18+00:00,2961,0.18541909661144018,-2453.944699,-3453.159474,-6395.339133,-2370.6189070667256,-3388.3557
2019-01-01 00:40:18+00:00,2961,0.1861135410144925,-2174.711811,-3235.544662,-6606.672588,-2088.6855656209477,-3168.34394
2019-01-01 00:41:18+00:00,2961,0.18680798541754484,-1886.692036,-3010.222074,-6797.026044,-1798.2160311580176,-2940.8606
```

```
epoch,kp
2019-01-01 00:00:00+00:00,1.0
2019-01-01 03:00:00+00:00,1.333
2019-01-01 06:00:00+00:00,2.667
2019-01-01 09:00:00+00:00,1.667
2019-01-01 12:00:00+00:00,2.0
2019-01-01 15:00:00+00:00,0.333
2019-01-01 18:00:00+00:00,0.0
2019-01-01 21:00:00+00:00,0.667
2019-01-02 00:00:00+00:00,0.0
2019-01-02 03:00:00+00:00,0.0
2019-01-02 06:00:00+00:00,0.0
2019-01-02 09:00:00+00:00,0.0
```

```
epoch,f107
2019-01-01 00:00:00+00:00,69.86666666666666
2019-01-02 00:00:00+00:00,72.39999999999999
2019-01-03 00:00:00+00:00,70.53333333333335
2019-01-04 00:00:00+00:00,69.56666666666666
2019-01-05 00:00:00+00:00,69.0
2019-01-06 00:00:00+00:00,69.76666666666667
2019-01-07 00:00:00+00:00,69.3
2019-01-08 00:00:00+00:00,68.86666666666667
2019-01-09 00:00:00+00:00,68.16666666666667
2019-01-10 00:00:00+00:00,67.53333333333333
2019-01-11 00:00:00+00:00,66.3
2019-01-12 00:00:00+00:00,67.23333333333333
```

Jan 20, 2026

Class #42: Dr. Garcia Meeting & Work Block

- General notes from Dr. Garcia lecture:
  - Methodology section due on February 19th
  - We will have individual check-ins, then a lesson on how to write the methodology section paper
    - LOOK AT LITERATURE to see how to format the methodology portion of my project
    - Software, databases, etc. all need to be CITED
  - I should send the methodology section paper around **Feb. 12th** or before so it can be read + reviewed by Dr. Garcia
  - Science Fair on March 2nd
  - Feb. Oral presentation day is Feb 23rd
    - I can present using my fully completed poster OR a slideshow -> a week before I can send my presentation to Dr. Garcia for edits + practicing
    - To get my feedback for my presentation, I should see Dr. Garcia at lunch or a spare block before the Science Fair (March 2nd)
  - The Science Fair poster can be in a flat poster format, or Tri-Fold
- Main working focus for today:
  - Continue building ML-ready ephemeris for TensorFlow -> merge propagations ephemeris with space weather ephemeris
  - Fix propagation inaccuracies by adding a leap second to the UTC time

Jan 8, 2026

Class #41: Mentor Meeting & STUDYING FOR MIDTERMS

- Meeting with mentor at 1:00 pm (see January Mentor Meeting Notes tab)
- Sent meeting summary email

Jan 6, 2026

Class #40: STUDYING FOR MIDTERMS

🗓️ February Daily Notes

DAILY NOTES - FEBRUARY

Feb 27, 2026

Class #55: {Oral Presentation #2 Day 3}

Work Block

- Main working focus for today:
    - Finalize February logbook + add/plan March calendar information (see March Calendar tab)
    - Add in my Results Data Collection/Test Results and Discussion/Conclusions data into my logbook tabs (see Results - Data Collection tab and Analysis, Conclusions, & Significance tab)

Feb 25, 2026

Class #54: {Oral Presentation #2 Day 2}

Watching Presentation x1 & Work Block

- WATCHED PEERS' ORAL PRESENTATIONS TODAY (see Major Assessments -> Oral Presentation #2 tab) -> notes/questions about other presentations
- Main working focus for today:
    - Finalizing February logbook details + planning March calendar (see March Calendar tab)
    - Find mentor meeting times for March and send an email to my mentor to confirm what times he wants to meet

Feb 23, 2026

Class #53: {Oral Presentation #2 Day 1}

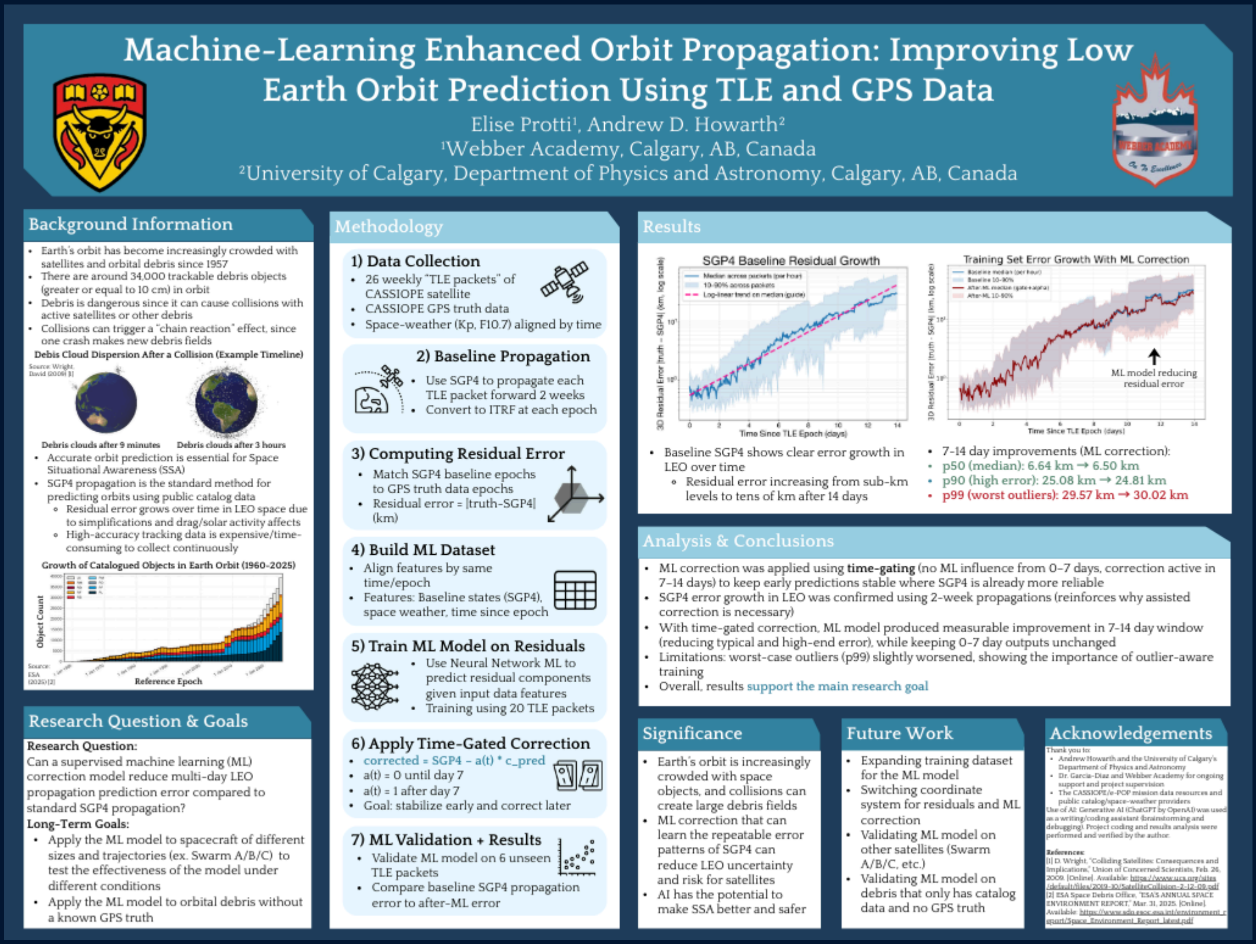
Presenting & Watching Presentations x3

- Presented by Oral Presentation #2
- WATCHED PEERS' ORAL PRESENTATIONS TODAY (see Major Assessments -> Oral Presentation #2 tab) -> notes/questions about other presentations

Feb 19, 2026

Class #52: Work Block & Mentor Meeting

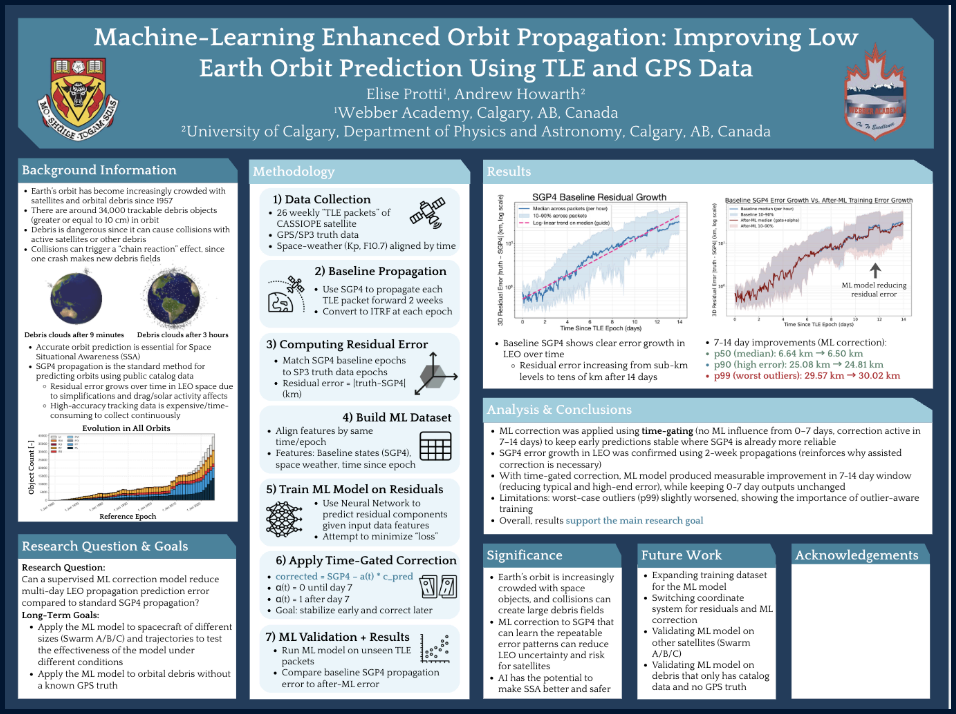
- Meeting with mentor at 10:00 am (see February Mentor Meeting Notes tab)
- Sent meeting summary email
- Main notes from Dr. Garcia lecture today:
  - Should not go off topic during oral presentation -> should rehearse a lot and be comfortable/informal but on-topic
  - Should emphasize that my project was not mentor-run, and I did the majority of the work -> show by my knowledge
  - Should answer questions clearly and show my knowledge + provide them with "deeper" knowledge
  - Lots of eye contact on the judge -> never turn my back on a judge when pointing things out on the poster
  - No que-cards and stay confident
  - Look over the judging rubric so that I make sure I have included/talked about all the necessary information
- Main working focus for today:
  - Finalize poster design and incorporate OP #2 edits from mentor + Dr. Garcia into the poster so it can be ready for printing over the weekend (see Major Assessments -> Oral Presentation #2 tab)
  - Add more info to CYSF portal



Feb 17, 2026

Class #51: Work Block & Dr. Garcia Meeting

- Main working focus for today:
  - Incorporate methodology section paper feedback into paper + references check (see Major Assessments -> Methodology Section Paper tab)
  - Update CYSF portal with up-to-date information so that I have all the categories filled in before the Science Fair

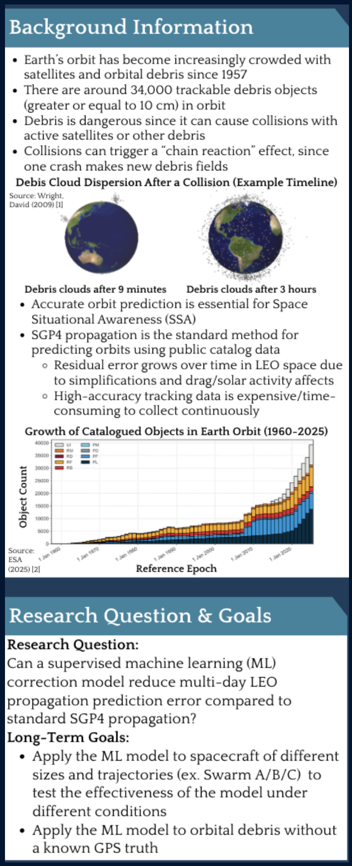


- Finalize poster design so I can print over the weekend (add acknowledgments section and ask Dr. Garcia about this) (see Major Assessments -> Oral Presentation #2 tab)

Feb 11, 2026

Class #50: Work Block & Mentor Meeting

- Meeting with mentor at 1:00 pm (see February Mentor Meeting Notes tab)
- Sent meeting summary email
- Main working focus for today:
  - Continue writing methodology section paper (see Major Assessments -> Methodology Section Paper tab)
    - Specific focus on the ML model process and the exact indicators we used to check its performance (p50, p90, p99)
  - Create background/research question section for poster (see Major Assessments -> Oral Presentation #2 tab)



- I pasted in and summarized some information about the background information of my project and the research questions
- Information taken from both my introduction section paper and my research proposal #1 slides (see Major Assessments tab)

Feb 9, 2026

Class #49: Work Block

- General notes from Dr. Garcia lecture today:
    - Only have a few weeks left in February to complete the project
        - Methodology draft due on the 12th -> methodology paper due on the 19th
        - Oral presentation on the 23rd
        - CYSF portal due on March 4th -> I should be copying and pasting elements into there whenever I have some time
- Main working focus for today:
    - Go over Intro section paper advice + comments for general writing/references comments that I can apply to my methodology paper writing
    - Continue writing methodology section paper
        - Specifically, work on citing data sources and how I converted them into tables (see Major Assessments -> Methodology Section Paper tab)

Feb 5, 2026

Class #48: Mentor Meeting & Dr. Garcia Meeting & Work Block

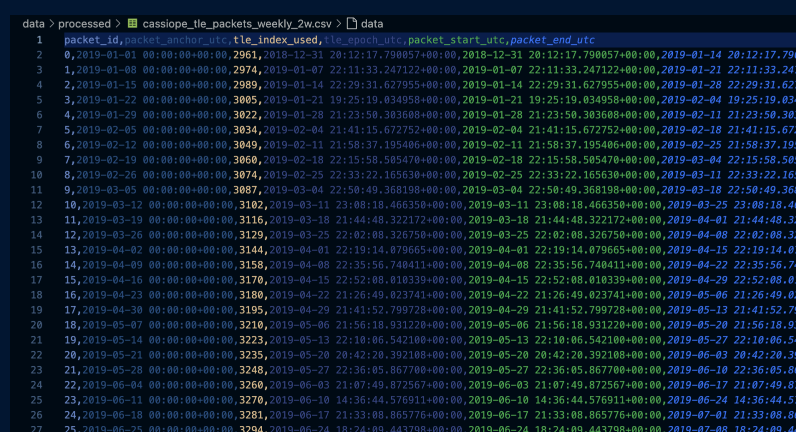
- Meeting with mentor at 9:00 am (see February Mentor Meeting Notes tab)
- Sent meeting summary email
- General notes from Dr. Garcia meeting today:
    - Science Fair judging
        - Gold = 100%
        - Silver = 95%
        - Bronze = 90%

        - HM = 85%
    - After the science fair, there will be 1 more oral presentation
- Main working focus for today:
    - Started listing ideas for methodology intro overview + material/tools used in bullet points then converted to professional text
    - Start official Methodology paper intro writing (see Major Assessments -> Methodology Section Paper tab)
        - Started working on the Data Collection and Organizing/Data Preparation, Parsing, and Cleaning sections

Feb 3, 2026

Class #47: Work Block

- General notes from Dr. Garcia lecture today:
    - Should be using my February calendar chunked with tasks so that all class blocks are filled
    - Should have tentative due dates in my calendar
    - A meeting in class with Dr. Garcia should only take about 10-20 minutes of time in February -> NOT THE WHOLE BLOCK
    - Dimensions of the Tri-Fold: 3x(80x150) cm
    - Dimensions of the Poster: 4ft x 3ft
- Main working focus for today:
    - Continued work on changing my dataset to tailor to 1 TLE propagated for a 2 week epoch to feed into my ML model



```
data > processed > cassiope_tle_packets_weekly_2w.csv > data
packet_id,packet_anchor_utc,tle_index_used,tle_epoch_utc,packet_start_utc,packet_end_utc
0,2019-01-01 00:00:00+00:00,2961,2018-12-31 20:12:17.790057+00:00,2018-12-31 20:12:17.790057+00:00,2019-01-14 20:12:17.7900
1,2019-01-08 00:00:00+00:00,2974,2019-01-07 22:11:33.247122+00:00,2019-01-07 22:11:33.247122+00:00,2019-01-21 22:11:33.2471
2,2019-01-15 00:00:00+00:00,2989,2019-01-14 22:29:31.627955+00:00,2019-01-14 22:29:31.627955+00:00,2019-01-28 22:29:31.6279
3,2019-01-22 00:00:00+00:00,3005,2019-01-21 19:25:19.034958+00:00,2019-01-21 19:25:19.034958+00:00,2019-02-04 19:25:19.0349
4,2019-01-29 00:00:00+00:00,3022,2019-01-28 21:23:50.303608+00:00,2019-01-28 21:23:50.303608+00:00,2019-02-11 21:23:50.3036
5,2019-02-05 00:00:00+00:00,3034,2019-02-04 21:41:15.672752+00:00,2019-02-04 21:41:15.672752+00:00,2019-02-18 21:41:15.6727
6,2019-02-12 00:00:00+00:00,3049,2019-02-11 21:58:37.195406+00:00,2019-02-11 21:58:37.195406+00:00,2019-02-25 21:58:37.1954
7,2019-02-19 00:00:00+00:00,3060,2019-02-18 22:15:58.505470+00:00,2019-02-18 22:15:58.505470+00:00,2019-03-04 22:15:58.5054
8,2019-02-26 00:00:00+00:00,3074,2019-02-25 22:33:22.165630+00:00,2019-02-25 22:33:22.165630+00:00,2019-03-11 22:33:22.1656
9,2019-03-05 00:00:00+00:00,3087,2019-03-04 22:50:49.368198+00:00,2019-03-04 22:50:49.368198+00:00,2019-03-18 22:50:49.3681
10,2019-03-12 00:00:00+00:00,3102,2019-03-11 23:08:18.466350+00:00,2019-03-11 23:08:18.466350+00:00,2019-03-25 23:08:18.466
11,2019-03-19 00:00:00+00:00,3116,2019-03-18 21:44:48.322172+00:00,2019-03-18 21:44:48.322172+00:00,2019-04-01 21:44:48.322
12,2019-03-26 00:00:00+00:00,3129,2019-03-25 22:02:08.326750+00:00,2019-03-25 22:02:08.326750+00:00,2019-04-08 22:02:08.326
13,2019-04-02 00:00:00+00:00,3144,2019-04-01 22:19:14.079665+00:00,2019-04-01 22:19:14.079665+00:00,2019-04-15 22:19:14.079
14,2019-04-09 00:00:00+00:00,3158,2019-04-08 22:35:56.740411+00:00,2019-04-08 22:35:56.740411+00:00,2019-04-22 22:35:56.740
15,2019-04-16 00:00:00+00:00,3170,2019-04-15 22:52:08.010339+00:00,2019-04-15 22:52:08.010339+00:00,2019-04-29 22:52:08.010
16,2019-04-23 00:00:00+00:00,3180,2019-04-22 21:26:49.023741+00:00,2019-04-22 21:26:49.023741+00:00,2019-05-06 21:26:49.023
17,2019-04-30 00:00:00+00:00,3195,2019-04-29 21:41:52.799728+00:00,2019-04-29 21:41:52.799728+00:00,2019-05-13 21:41:52.799
18,2019-05-07 00:00:00+00:00,3210,2019-05-06 21:56:18.931220+00:00,2019-05-06 21:56:18.931220+00:00,2019-05-20 21:56:18.931
19,2019-05-14 00:00:00+00:00,3223,2019-05-13 22:10:06.542100+00:00,2019-05-13 22:10:06.542100+00:00,2019-05-27 22:10:06.542
20,2019-05-21 00:00:00+00:00,3235,2019-05-20 20:42:20.392108+00:00,2019-05-20 20:42:20.392108+00:00,2019-06-03 20:42:20.392
21,2019-05-28 00:00:00+00:00,3248,2019-05-27 22:36:05.867700+00:00,2019-05-27 22:36:05.867700+00:00,2019-06-10 22:36:05.867
22,2019-06-04 00:00:00+00:00,3260,2019-06-03 21:07:49.872567+00:00,2019-06-03 21:07:49.872567+00:00,2019-06-17 21:07:49.872
23,2019-06-11 00:00:00+00:00,3270,2019-06-10 14:36:44.576911+00:00,2019-06-10 14:36:44.576911+00:00,2019-06-24 14:36:44.576
24,2019-06-18 00:00:00+00:00,3281,2019-06-17 21:33:08.865776+00:00,2019-06-17 21:33:08.865776+00:00,2019-07-01 21:33:08.865
25,2019-06-25 00:00:00+00:00,3294,2019-06-24 18:24:09.443798+00:00,2019-06-24 18:24:09.443798+00:00,2019-07-08 18:24:09.443
```

        - Checked for any gaps in my GPS data for the time period, and wrote code to skip rows that had spotty data, so that the propagated “TLE packets” would be accurate for my ML model
        - My TLE packets weekly dataset for all 26 data points being used is above
    - Building methodology paper skeleton (create headings and consolidate all sections needed) (see Major Assessments -> Methodology Section Paper tab)

- Example sections that I included in the skeleton structure:
    - Data Collection and Organization
    - Data Preparation, Parsing, and Cleaning
    - Baseline Orbit Propagation and Residual Calculations
    - Machine-Learning Correction
    - Machine-Learning Evaluation and Reporting
    - Evaluation of Results and Statistical Analysis

🗓️ March Daily Notes

DAILY NOTES - MARCH

Mar 3, 2026

Class #56:

🗓️ April Daily Notes

DAILY NOTES - APRIL

Apr 1, 2026

Class #:

🧑‍💻 Mentor Meeting Notes

MENTOR MEETING NOTES

See individual months.

Mentor Meeting Room: https://ucalgary.zoom.us/my/ahowarth

Mentor Meeting Availability:

- **Flexible days:**
  - **Monday mornings**
  - **Wednesdays**
  - **Thursdays**
  - **Fridays (after 9 am)**
- Busy days:
  - Tuesdays
  - Monday afternoons

🗓️ September Mentor Notes

MENTOR MEETING NOTES - SEPTEMBER

Sep 25, 2025 4:30 PM

Meeting #4: Finalizing Project Idea + Discussing Software

Agenda:

- What citation style should I use? (APA 7th edition - common for astronomy and astrophysics, IEEE - common for engineering, in most research papers I am using)
- What project ideas did you come up with?
- Share my specific project idea that I have found + document with other potential options
  - **TLE-only ML model that predicts short-term collision risk** (flags pairs of objects likely to have a close approach soon), predict future collisions? **FOCUS ON QUICK, EFFICIENT DATA RETRIEVAL OF A SPECIFIC SATELLITE/PIECE OF DEBRIS**
    - Risk score (0-1) OR alert/no-alert signal?
  - ML models can learn "danger patterns" -> **small altitude gap/ high closing speed/converging planes /increase in uncertainty**
- Show flow chart -> is this what the simplified steps are?
- Where would we put TLE data? Excel? Python libraries?
  - What programs would we use (Visual Studious, TensorFlow)?
  - Should I list a bunch of options in my R.P. then finalize it later in the project?

Notes:

- I should use a "Geophysical Research Letters (GRL)" -> AGU version or IEEE citation style

- Mentor project ideas:
  - **Use AI to improve upon orbital prediction of satellites with TLE data -> orbit propagation -> can compare GPS data with predicted TLE data -> use AI to take historical data and feed it into ML model -> compare accuracy -> can predict collisions**
  - CASSIOPE has GPS data
  - SWARM satellites (3 satellites) -> GPS data is 100% available
  - Use satellites with different orbits to see how the ML model works
- Problems with my idea:
  - Need to propagate a lot of data to see if anything will collide with a satellite/spacecraft
  - Might be too complex for a year-long investigation
- CASSIOPE/Swarm -> have raw GPS data and USEFUL: **orbit determination**
  - GPS data that is converted to regularly spaced data points that is easy for analysis, **orbit with uncertainty of 5 meters**
- Use Python for inputting TLE and GPS data -> pandas for scientific dataframe for information > can do calculations on it as well
- TensorFlow/Pytorch:
  - A program for implementing machine learning model

Tasks to Complete:

- ~~Research more about TensorFlow and Pytorch -> due for next meeting (Oct. 3)~~
- ~~Finish R.P. and send to mentor for comments -> due end of Sep./beginning of Oct.~~
- ~~Send October meeting time possibilities -> due Sep. 25~~

Sep 17, 2025 10:00 AM

Meeting #3: Data + Development Environment Discussion

Agenda:

- What performance metrics will we be measuring? What data will come from every test that we run?
- Confirm how we are gathering data -> open source database, which one? images from where?
- Concrete steps to starting research proposal

- Tell me logistically what you think will happen in the project since I am having trouble wrapping my head around the processes -> what programs are we using? What interface? What tests will be run? What results will be gathered?
- I want to set up my "working environment" with Python libraries of data to work on
- Does UofC have educational access to any data sources?
- Is there a specific niche that you would like to pursue?
- Looked at many data sources:
    - Space-track.org -> TLE data of satellites and space debris (how do I access, API needed?)
    - CelesTrack.org -> TLE data of orbital debris, smaller data set with less historical data
    - Kaggle.com -> CVS data (data separated by commas, easy to import-export)
        - https://www.kaggle.com/datasets/atharvasoundankar/global-space-exploration-dataset-2000-2025
    - Telescope Observational Data?
        - CASSIOPE plasma wave detection from space debris
    - Radar Observational Data?
        - U.S. Space Surveillance Network (SSN)
    - LeoLabs Data
        - Large commercial catalog of data + high accuracy -> integration with AI

Notes:

- Radars/optical data to make TLE data -> gives orbits of each item in catalog
- Should do more investigation into LeoLabs -> LeoLabs identifies objects with data given
- Look at optical space data from satellites in space may be more effective (small objects might not be able to be tracked from the ground)
- Optical image recognition -> measured metrics: (training data set 80% -> apply model to it, over 20% of the data set is used to look at results)
- **STK software** used to enter TLE and show the resulting orbit (so I do not have to do manual work to find the orbit of the debris) -> **paid version**
- Using optical data is challenging because:
    - Spacecraft is spinning -> hard to teach computer to have image-pattern recognition -> need relative motion/sequence of frames to analyze what is orbital debris vs. what is a star
- Tools used at the UofC: **VS code** development environment (Copilot built in)

- **TensorFlow** can be used for optical data + more?
- Project idea: can answer previously solved issues with TLE data and AI
- Project idea: quick answers about TLE data for specific objects in space (ex. satellites) given a time and a location
- Need to find project idea -> then I can thing about variables and procedures
- TLE software needed to analyze TLEs to check machine learning algorithm accuracy

Tasks to Complete:
- ~~Research different softwares to use for development environment -> due Sep. 25~~
- ~~If using TLE data, need to think about unique project idea - look at conclusion paragraphs of related projects in AI/ML and orbital debris -> due Sep. 25~~
- ~~Look into LeoLabs to get extra information about them -> due Sep. 25~~

Sep 11, 2025 12:30 PM

Meeting #2: Intro to Research Proposal

Agenda:
- Ask for more research papers/research sources to use for the R.P.
- How much research do I need for my R.P.? -> what specific topics?
- Confirm how we are gathering data -> open source database, which one? images from where?
- Confirm materials needed to complete the project
- R.P. deadline -> end of Sept. (no later than Oct. 9) for mentor review

Notes:
- As many research papers as possible, around 6-8 references in the intro of the R.P.
    - References to different methods for the proposal
- Scientific problem = space debris (what others have proposed) and how I propose
- Devote time to learning about orbital debris, machine learning
- Space-track.org, provides 2-line elements
    - 2 lines of data, gives info about the orbit of a satellite at a given time (has all US space command satellites -> starlink, geosynch, scientific, etc.), gives a TLE -> given the right software, can find the orbit of the object, LEO debris + satellites less accurate than far away objects

- Optical images may be easier for a dataset -> where machine learning started (pattern recognition, etc.)
- Good training data with accurate data is essential
    - Lots of pre-made algorithms to make small adjustments to the project
    - Train model to use sequences of images to determine objects

Tasks to Complete:

- ~~Send current progress on R.P. at around Sept. 29th before his trip~~
- ~~Send meeting follow up email~~
- ~~Research more about potential data sets, machine learning algorithms, and space debris generally~~

Sep 3, 2025

Meeting #1: Summer Catch Up + Questions

Agenda:

- Explain classroom guidelines
- Explain monthly calendar
- Catch up on summer work
- What type of project is mine? **Innovation**, Experimental, Study
- Ask about future meeting times (how frequent? in person? during ASP classes -> send schedule)
    - Once a week at the start
    - Zoom calls until other notice
    - Will try to call during ASP class times
- Mention school extracurriculars + time commitments
    - Volleyball, WASSU, Legacy
- Confirm which project we are doing (classifying orbital debris for AI system)
- *Optional: research proposal guidelines*

Notes:

- Completed agenda + went over research proposal guidelines
- Mentor works on Tuesdays and Fridays at the UofC, other days he works at home
- Prefers Zoom calls
- **Flexible days:**

    - **Monday mornings**
    - **Wednesdays**
    - **Thursdays**
    - **Fridays (after 9 am)**
- Busy days:
    - Tuesdays
    - Monday afternoons
- Confirmed we are working on the orbital debris project
- Locating orbital debris from images is difficult -> don’t worry if this is troublesome, may use data sets instead

Tasks to Complete:

- ~~Send research proposal guidelines to mentor -> due Sep. 3~~
- ~~Send potential meeting dates for September to mentor -> due Sep. 3~~

🗓️ October Mentor Notes

MENTOR MEETING NOTES - OCTOBER

Oct 30, 2025 10:00 AM

Meeting #8: Oral Presentation Feedback

Agenda:
- Show mentor my oral presentation slides so far, and ask him to give me any comments/suggestions he thinks I should add

Notes:
- I sent him the presentation link -> he will have feedback for me by tomorrow (Oct. 31st) afternoon

Tasks to Complete:
- ~~Finish the oral presentation + edit it using the comments that he will provide tomorrow (Oct. 31st) and present my presentation on November 3rd~~

Oct 22, 2025 10:00 AM

Meeting #7: Comments about R.P.

Agenda:
- Discuss general + specific comments about my research proposal (discuss any edits that need to be made before I submit my proposal tonight)
- Talk about next step in the project -> oral presentation

Notes:
- Edits for me to update on my research proposal:

    - If I have time, I can add space weather inputs to the flowchart in my methodology, but this may be too specific/not important enough to mention in my road methodology
    - I should mention/emphasize more than space weather indicators (geomagnetic weather data -> Kp past and predicted data, solar activity data -> F10.7 data)
- Oral presentation new presentation date -> November 3, 2025
    - Should focus on making the data on the slides readable and simple so others can easily understand everything (especially since I have a complicated methodology) -> should use FLOWCHART on slides

Tasks to Complete:
- ~~Complete basic rough outline of all sides of OP by next meeting -> due Oct. 30, presentation on Nov. 3~~

Oct 16, 2025 1:00 PM

Meeting #6: Comments/Clarification about R.P.

Agenda:
- Discuss general + specific comments about my research proposal
- Next steps after the R.P. is completed? -> how should I tentatively plan my November?

Notes:
- Should do some quick research on the SGP4 propagation process to add to my proposal -> takes into account Earth's shape, drag, radiation from the sun, and gravitation effects from the Sun and the Moon
- Data acquisition is very quick -> just need to go to open source sites and download data
    - TLE and GPS data for satellites -> CASSIOPE (elliptical orbit), Swarm A/C (lower pair - always travel together), B (higher elevation)
- Research generally on CASSIOPE and Swarm A/B/C orbits
- After data acquisition, need to write Python code to open the data files
- Then, get acquainted with development environment that I will be using for the algorithm (TensorFlow or PyTorch)

        - Input data into model, output useful parameters

Tasks to Complete:

- ~~Finish research proposal -> due Oct. 22~~
- ~~Research more about TensorFlow and Pytorch -> due for next meeting (Oct. 3)~~

Oct 3, 2025 10:00 AM

Meeting #5: Clarification of Project Details

Agenda:

- Discuss what variables we will be using for the project
    - Variables (20 marks): The proposal has to establish the key variables of the study.
        - a. Independent variable: manipulated or treated in the study
        - b. Dependent variable: responding variable (effect, result, outcome)
        - c. Controlled variable: kept constant during the study
        - d. Confounding variable: May influence the effect of independent variables
- What ML approach will we use?
    - Supervised learning, reinforcement learning, or another approach? -> I think supervised learning is the best since we can teach the model through giving the model known inputs (GPS data)
    - Will we use neural networks or traditional regression-based models? -> will we be using more complicated neural networks or just a data-driven traditional regression model (will we have to make these or can we get these online?)
- Does this project already exist? What is innovative about it?
- What other advice do you have for me writing the research proposal?

Notes:

- Leaving tomorrow afternoon for Europe -> back next Friday afternoon
- How broadly/briefly should I mention the variables for the project -> can list a few potentials + add more over time if necessary
- Expecting we will use Convolutional Neural Networks (CNN) + supervised machine learning -> model can change after we get data

- Mentor thinks that our project is innovative and already niche
- Advice for research proposal:
  - Should avoid writing a proposal like a diary -> can use a 3rd person POV if you want
  - Should make the writing professional (and not give unnecessary information about intermediary steps)
    - DON'T GET BOGGED DOWN WITH TINY DETAILS
- **After the research proposal and the oral presentation, we would start data gathering and analysis**

Tasks to Complete:

- ~~Finish research proposal -> due Oct. 22~~

📆 November Mentor Notes

MENTOR MEETING NOTES - NOVEMBER

Nov 19, 2025 10:00 AM

Meeting #11: Long Range Plan

Agenda:

- Talk about a specific long-range plan of tasks in the remaining 9/10 weeks of the project before the Science Fair (CYSF)
- Update about progress on Python course
    - I have practiced inputting data files into VS code and producing different Pandas Dataframes with them
- Go through background notes on SGP4 algorithms + decide on the one we will be using
    - When propagating orbits for CASSIOPE (and Swarm), what time period in advance do we propagate, and do we propagate an orbit for every single TLE data point we have? -> is there 1 TLE data point per day?
- How are we going to calculate residuals?

Notes:

- Discussed Skyfield and SPG4 packed library integrated into Python
    - Should check when these systems were last developed (recent?) and how many features they have
- Once I have SGP4 implemented, the position will be determined the future -> residuals should go up as time goes on
    - **Every minute** should output a time and position
    - Should look at coordinates - lat, long, altitude -> convert to x, y, z plane (geographic frame of reference)
        - I should look for what the SPG4 outputs in terms of time
        - TLE data is in the form of "classical orbital elements", but the SGP4 propagation system will probably take it in still

    - GPS data usually gives values every second, should pull out the times for every minute
    - We should predict the orbit in LEO for the CASSIOPE TLE data using the SGP4 algorithm for **2 weeks in advance**
    - SGP4, if it works, will be very quick -> if I have to change formats or doesn't work, this could be a challenge area
- Tensorflow - unknown data format, we need to research how to add the dataset into the algorithm, can maybe use a NumPy array or a CSV file?
- Datasets are the same on Swarm A/B/C as in CASSIOPE
    - Should write the code in the mindset that I may use different missions/data, should not overfix to just CASSIOPE
- **LONG RANGE PLAN**

1) Week 1 - SGP4 propagation and calculating residuals
2) Week 2- SGP4 propagation and calculating residuals
3) Week 3 - TENSORFLOW - input data
4) Week 4 - TENSORFLOW- running supervised machine learning training set
5) Week 5 - TENSORFLOW - outputs in Tensorflow
6) Week 6 - TENSORFLOW - compare outputs with GPS data for accuracy predictions
7) Week 7 - Tentative: Redo process with a Swarm satellite of choice
8) Week 8 - Analysis of data results and discussions of error + conclusions

Tasks to Complete:

- ~~Practice importing the data sets (our downloaded data for the project or the datasets in the Python course for practice) to create a Pandas dataframe -> due Dec. 3~~
- ~~Extended background research into SGP4 ALGORITHM OF CHOICE and how to import it into Python, Pandas, TensorFlow, etc. - due Dec. 3~~
- Input all the acquired data into a Python DataFrame -> due Dec. 7
- Start propagating orbits for 2 weeks in advance -> need to pay attention to what coordinate systems I need for the data -> due Dec. 7
- (Optional) Start calculating residuals process
- (Optional) Start research into Tensorflow, and what format I will need to input the input data set in

Nov 13, 2025 12:30 PM

Meeting #10: Getting Comfortable with Python and Pandas + Next Steps

Agenda:
- Inform about potentially missing the meeting on Thursday, November 27 (because of November provincials)
- Update about my progress on the “Data Analysis Using Python” course
- Report back my progress on the download of Kp, F10.7, TLE, and GPS data
- Report back my progress of setting up VS Code and Anaconda
  - Should I do some practice things/projects on VS Code to get more comfortable with it?
- What are the next steps of the project?
  - Should I continue progressing through the Data Analysis Python course?
  - Watching related Python/Pandas videos?
  - Should I research the SGP4 algorithm and how to use it? Videos on SGP4?
- What tasks should I complete before our next meeting?

Notes:
- The next step of the project is to get familiar with VS Code and Python, practice reading the CSV or txt data into Pandas using a Pandas Dataframe, long-run: using tensorflow to run the data through a machine learning algorithm
  - Can calculate sums, averages, query(), and plot points
  - One of the best things to do to become a good, efficient coder is to learn about the capabilities of the library/package that you are using
- Txt files and CSV files are very similar and work well with Pandas
  - The read_csv() command in Python can read ANY type of text file (CSV or txt)
    - sep = “,”
    - names -> can assign names to each column
    - skiprow or skipline commands to skip headers/context at the beginning

Tasks to Complete:

- ~~Continue working through the Python course~~
- ~~Practice importing the data sets (our downloaded data for the project or the datasets in the Python course for practice) to create a Pandas dataframe -> due Nov. 19~~
- ~~(Optional) Background research into Pandas + Dataframes, SGP4 propagation systems -> due Nov. 19~~
    - ~~See if there are any free versions that I can use~~

Nov 5, 2025 9:00 AM

Meeting #9: Data Acquisition Steps + Path Forward

Agenda:

- Discuss the next steps of the project
- Can you give me a list of steps of how I should do the data acquisition?
- How should I download the TLE and GPS data -> Space-track.org?
    - Should I just download all the data onto my laptop?
- What time period/epoch should I download the data for?
- Should I do some research on how to use Python and Pandas for data analysis

Notes:

- Should get CASSIOPE data and Swarm A/B/C data on Space-track.org
    - Can start with CASSIOPE data
    - Should start with 6 months or 1 year worth of data (2019) -> should be around 1.5 gigabytes (SP3s - GPS data)
    - CASSIOPE has some holes in GPS data, whereas Swarm has 100% coverage
- GET TLE DATA FOR CASSIOPE: Go to ELSET Search -> Search for CASSIOPE -> select 6 month or 1 year time range -> Download “As TLE”
- GET GPS DATA FOR CASSIOPE: Go to UCalgary e-POP Data website -> CASSIOPE Orbit Ephemeris SP3 -> set time range -> download data
- F10.7 is a measure of the signal from the sun at 10.7cm wavelengths -> follows the 11-year solar cycle (related to number of sunspots) - 1 value per

day, can get this measurement from Natural Resources Canada, NOAA, LASP, etc.

- Kp measures magnetic disturbance level (magnetic activity level) on the Earth, high Kp = lots of energy put into the atmosphere, and atmosphere expands (affects orbits of spacecraft in LEO)
- Should enter the data into a text file (txt) -> can save as a file on my desktop
- Pandas is quite easy to use -> can "read" a text file separated by spaces or commas, returns a data object with all desired data
    - This data object can have operations performed on it as well
    - Can use VS code (has Copilot built in), should also search for alternatives
    - Python is the main language, Anaconda (can be easily downloaded) is a python package that has specific data packages built into it -> it has Pandas, etc.

Tasks to Complete:

- ~~Finish downloading the GPS (SP3) and TLE data for the CASSIOPE satellite for a 6 month time period in 2019 -> due Nov. 13~~
- ~~(Optional) Research more about the space weather indicators and start the download process on them -> due Nov. 13~~
- ~~(Optional) Background research on development environments and Python packages (VS code, Anaconda, etc.) -> due Nov. 13~~

📅 December Mentor Notes

MENTOR MEETING NOTES - DECEMBER

Dec 17, 2025 12:30 PM

Meeting #14: Winter Break Plan

Agenda:

- I need to clearly outline what work I will be doing over the Winter Break
- Plan out December and January clearly
- Meeting plan over the Break and in January after midterms?

Notes:

- On January 8, January 22/23 we will have a mentor meetings and we will discuss then when our future meetings will be + our progress on TensorFlow
- Over Winter Break:
    - I should start researching TensorFlow, but we should begin this concrete process when we both get back in January
    - I should finish the SGP4 propagations (converting to the same coordinate frame) and preparing the input data set
- Mentor will be on a trip from January 26-30, so we cannot have a meeting during this period of time

Tasks to Complete:

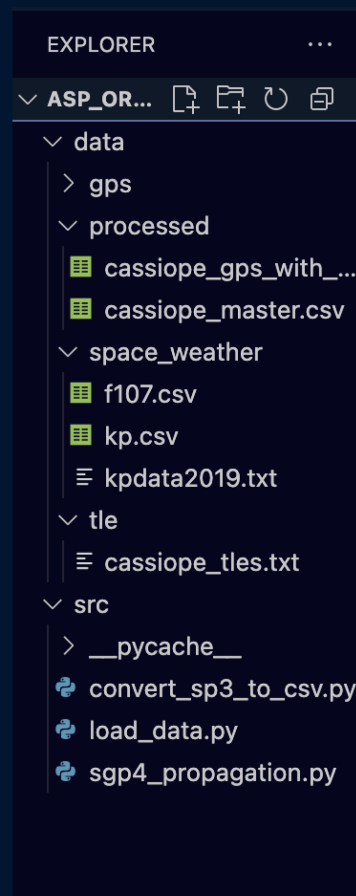
- ~~Finish SGP4 propagations by converting to a uniform coordinate system~~
- ~~Finish building input dataset for TensorFlow~~
- ~~Extended research into TensorFlow so we can be ready to start the TensorFlow ML process in January and into February~~

Dec 11, 2025 9:00 AM

Meeting #13: VS Code Update

Agenda:

- Go over what I have done since the last meeting:
  - Made a suitable environment in VS Code within the "ASP_ORBIT_ML" folder on my computer
  - Go through the folder + subfolder names in VS code + their functions --------->
  - Go over the output csv documents I have made
  - Show my mentor my SGP4 propagated output and ask if this is what we are looking for?
    - How many TLEs should we use? Should we propagate them all in advance?
- My data types are in 2 different measurement/coordinate systems:
  - SP3/ GPS truth data is in **geographic/ITRF Cartesian frame**, (Earth fixed, rotating with Earth)
  - SGP4 propagated states are in **Earth-centred inertial** frame (TEME/ECI), (non-rotating with respect to distant stars)
  - I put the error value I initially got for my first propagated TLE (to the first epoch/time period given by my SP3/GPS truth data) into ChatGPT to troubleshoot -> this error number was around ***13471***
    - ***THIS POINTS TO DIFFERENCES IN COORDINATE SYSTEMS***



- Steps that I completed on VS code/my computer files since our last meeting:
  - Set up asp_orbit_ml project folder with different data/ and src/ folders in it
  - Saved the CASSIOPE TLEs as a txt file
  - Unzipped 137 SP3 files into the sp3_raw folder
  - Converted the SP3 files all to one .csv file (parsing through them in VS code)

        - I verified these SP3 positions are in km and make sense in context
    - Cleaned up the Kp text file in Google Sheets by building epoch + Kp columns, and I exposed this as a .csv file
    - I also changed the format of my F10.7 data to a .csv file
    - Updated a file called load_data to load the TLE, GPS, Kp and F10.7 (from Julian date to normal datetime) data -> converted the times all to regular calendar datetimes + built a combined master table
        - Note: SGP4 library requires dates to all be in Julian Date, so I convert them to that before feeding it into the SGP4 system
    - Generated a processed master .csv file with GPS truth + Kp + F10.7 aligned by epoch
    - Installed the sgp4 Python library into my Anaconda environment in VS code -> created a sgp4_propagation.csv file to run the SGP4 using the first CASSIOPE TLE and propagated at every GPS epoch that is in my dataset
    - Observed a very large position_error_km at all the epochs -> indicated a coordinate-frame mismatch
        - SP3 in ITRF
        - SGP4 in TEME
- One problem I have: lots of NaN values in the space-weather columns (Kp and F10.7)
    - This happens because the data rows are only joined if the timestamps are exactly IDENTICAL (slight discrepancies between SP3 (GPS truth), Kp, and F10.7 times)
        - SP3 times are about every second
        - Kp is about every 3 hours
        - F10.7 is once a day at times almost at the hour
    - Solution to this: I need to match to the closest time/same day and use that Kp/F10.7 value -> NEED TO WORK ON THIS
- How should I send the TLE with its propagated state to you?

Notes:
- Hard to convert TEME to ITRF coordinate system
- Do not need extreme accuracy for this project -> very complex to accurately convert between the two
- Same difficulty to convert between the 2 types (inverting the transformation matrix)
- Root of sum of squares gives 3D position error

    - sqrt((x-x1)^2 + (y-y1)^2 + (z-z1)^2) -> EUCLIDEAN DISTANCE FORMULA
- I should export the TLE data with the SGP4 propagated states to my mentor
    - Just send the files
- Mentor thinks the TensorFlow process will be difficult, so I need to get my dataset together as fast and accurate as possible
- Mentor expects the position error for a satellite to be a few 10s of km after a day (less than 50km after a day)

Tasks to Complete:
- ~~Send mentor the formula I used for calculating 3D position error~~
- ~~Finish propagating the orbit of the satellite (1 TLE) for 2 weeks in advance, save this into a file, and send it to mentor~~
- ~~Get rid of the NaN values in my master DataFrame by aligning to the nearest hour/day (we do not need extreme accuracy)~~
- ~~Finish TLE propagations~~
- ~~(Optional) Research into the specifications of the TensorFlow input data set (what format, data types, etc.)~~

Dec 3, 2025 8:30 AM

Meeting #12: SGP4 Propagation Specifics

Agenda:
- Discuss the coordinate systems I should be using in the SGP4 propagation process
- What SGP4 algorithm should I use for the propagation process? -> have you done any more research into this?
- Have you done any more background research into TensorFlow?
- Should I propagate a TLE then you propagate the same one to compare the accuracy of the propagation system?
    - How far in advance should I propagate the data?
- What types of data are TLEs and GPS data?

Notes:
- Should use the most python-centered SGP4 programs (Python 3) -> usually the most helpful, modern, and have lots of active forums
- Need to decide the COORDINATE system we want to propagate in

    - See when the default coordinate system is for outputting on the SGP4 algorithm!
    - May have an option to choose a coordinate system to use -> the "best" one for me to use for this project is geographic (x, y, z) - ITRF or classical orbital elements (the form that a TLE is in), or J2000 (x, y, z), long/lat/altitude/ - geodetic
        - intract, crosstract, radial for residual calculations
- To start, take 1 TLE and propagate it for 2 weeks then send this to my mentor to compare my propagation to my mentor's SGP4 propagated state
    - GPS data is in geographic(xyz)/ITRF
    - TLE data is in the classical orbital elements - probably does not need to change the coordinate system for TLEs (can probably feed right into the SGP4 propagation system)

Tasks to Complete:

- ~~Finish importing our downloaded datasets into Python via Pandas -> due Dec. 7~~
- ~~Finish propagating orbits for 2 weeks in advance -> need to pay attention to what coordinate systems I need for the data -> due Dec. 7~~
- ~~Send my mentor my propagation results and the TLE I used so we can compare propagation accuracies between our SGP4 propagator systems -> due Dec. 11~~
- ~~Start calculating residuals process~~
- ~~(Optional) Start research into Tensorflow, and what format I will need to input the input data set in~~

🗓️ January Mentor Notes

MENTOR MEETING NOTES - JANUARY

Jan 22, 2026 10:00 AM

Meeting #16: ML Specifics

Agenda:

- Discuss my progress of finishing the input data set for the ML model
- What are the next steps forward?
    - What indicators are we going to use to measure the ML accuracy?
    - What is the ML model improving on -> residual error in x, y, z?
    - What is the “loss” function for the ML model and is it applicable?
- Discuss the future timeline for the project (TensorFlow time and how long the analysis might take)

Notes:

- Next steps for the project:
    - If you take a given TLE and propagate it for a specified window (ex. 2 weeks)
    - Should pick some data points with high/low Kp and high/low F10.7
        - Compare windows of low Kp vs one with a spike in Kp
        - Kp is the most important and will probably have the largest impact on the orbits of space objects since it has shorter cycles than F10.7
        - F10.7 data cycles every 11 years -> 6 months of data may not feature large changes in F10.7 data
    - SWARM data is very continuous and in a similar format to TLE for a next stage for the project potentially
    - Take 1 TLE (1 data point) -> start at its epoch and propagate it for 2 weeks -> can see residual error growth

        - The more of these trials I can do, the better until DIMINISHING RETURNS
        - Should make the cope so that it LOOPS for all TLEs to increase accuracy
        - 80% of “data packets” are used to train models, and 20% for testing -> does the ML model make the residuals better or worse?
            - Just give the ML model the TLE, Kp, and F10.7 and it will give it a propagated orbit for that TLE point -> compare with the real GPS truth to calculate the residuals after
        - The “loss function" for the ML model will be based on the residuals (want the “loss” to be 0) -> may not even need a loss function if I can just compare the residuals before and after (root, mean, square)
        - The ML model needs to be able to propagate the orbits in advance to predict the future orbits -> look for what the model is able to adjust in the SGP4 propagator to give a more accurate future position

Tasks to Complete:

- ~~Run the ML model using at least one “data package”/TLE~~
- ~~Work on making code in Python in VS code that can take a TLE and propagate it in a continuous loop in the ML model (more research into whether there are changeable parameters in SGP4 propagation?)~~

Jan 8, 2026 1:00 PM

Meeting #15: After Winter Break Check In

Agenda:

- Discussed my winter break progress and what I have done over break
    - My work on my input data set and my visual studio Python code
- Plan a meeting time after Midterms to meet again

Notes:

- I should start working with TensorFlow and use practice non-image data sets

- The analysis portion of the project will take some time (2+ weeks maybe?) so I should try to finish the project as soon as possible to get results by end of Feb for science fair

Tasks to Complete:

- ~~Finish SGP4 propagations by converting to a uniform coordinate system~~
- ~~Finish building input dataset for TensorFlow~~
- ~~Extended research into TensorFlow so we can be ready to start the TensorFlow ML process in January and into February~~

🗓️ February Mentor Notes

MENTOR MEETING NOTES - FEBRUARY

Feb 19, 2026 10:00 AM

Meeting #19: New Due Dates and Poster Coming Soon

Agenda:

- Get my mentor up-to-date with the upcoming deadlines in my ASP class and the Science Fair
    - Methodology section paper deadline is delayed until after the Science Fair, and it is now due on March 9th
    - My oral presentation is on Monday, February 23rd
    - The Science Fair is on March 2nd
- Discuss when I should send my finalized poster design to my mentor -> Friday, February 20th?

Notes:

- Future Work potential: switch coordinate systems from cartesian x/y/z to along track/across track/radial coordinate system (ML may perform better)
    - Reduces 3D residuals to a 1D residual (along track)

Tasks to Complete:

- Finish methodology paper draft #2 and send it to mentor and Dr. Garcia for edit suggestions (it will still be a rough draft) -> finish by March 2nd/3rd
- ~~Send my final poster design -> due by Friday, February 20th~~

Feb 11, 2026 1:00 PM

Meeting #18: Discussion of Graphs and Starting Poster Design

Agenda:

- Show final ML graph -> any feedback to improve or is this a good image for my poster?
- Do you have any preferred poster formats?
- Can I send you my methodology paper over the next few days for you to edit?

Notes:

- Should generate 2 plots (1 plot of ML training outputs and 1 plot of ML validation outputs)
- Mentor will send some posters examples in PowerPoint -> I can use Canva as well to complete my poster
- Mentor will send his edits to my methodology over the long weekend/early next week before its due (Feb. 19th)

Tasks to Complete:

- ~~Finish 2 ML plots + send to mentor -> finish by Feb. 13th~~
- ~~Finish methodology paper + send to mentor for edits -> finish by Feb. 12th~~
- ~~Finish oral presentation poster draft -> finish by Feb. 16th~~

Feb 5, 2026 9:00 AM

Meeting #17: ML Specifics

Agenda:

- What ways can we show the results of the ML and the residual error in the dataset on the poster?
- How will the ML training work -> is a 20/6 split okay?
- Give residual numbers -> ask if okay

Notes:

- 20/6 TLE training set should work well
- Most datasets have small data gaps so it is important to work around them
- Should do a consistency check of the residual data -> plot this to visually see exponential.curved growth of
- Should have these images on my poster:
    - Error growth with and without ML model (probably a scatter plot with a line of best fit) -> average residuals per day for training set, and another plot for testing set?
    - Can plot ML metrics -> **loss function**?, something to do with space weather indices
    - Methodology flowchart
    - Graphs from external sources of population growth of orbital debris in LEO

Tasks to Complete:

- ~~Try to plot some of the residuals so I can visually see the exponential growth of residual error values as time goes on -> tentatively finish by the end of the weekend (Feb. 8th)~~
- ~~Start ML TensorFlow process -> tentatively finish by the end of the weekend (Feb. 8th)~~
- ~~Start drafting methodology -> finish by Feb. 12th~~

🗓️ March Mentor Notes

MENTOR MEETING NOTES - MARCH

Mar 19, 2026 12:30 PM

Meeting #22:

Agenda:

-

Notes:

-

Tasks to Complete:

Mar 11, 2026 12:30 PM

Meeting #21:

Agenda:

-

Notes:

-

Tasks to Complete:

Mar 5, 2026 9:00 AM

Meeting #20:

Agenda:

-

Notes:

-

Tasks to Complete:

-

📅 April Mentor Notes

MENTOR MEETING NOTES - MARCH

Feb 25, 2026 9:00 AM

Meeting #:

Background Research

BACKGROUND RESEARCH

See individual months.

FUTURE RESEARCH SOURCES:
https://flypix.ai/blog/space-debris-mapping/
https://platform.leolabs.space/visualization
https://leolabs.space/
https://leolabs.space/ai/
https://www.esa.int/Applications/Observing_the_Earth/FutureEO/Swarm

🗓️ September Background Research

BACKGROUND RESEARCH - SEPTEMBER

Source 6: "A Machine Learning-Based Approach for Improved Orbit Predictions of LEO Space Debris With Sparse Tracking Data From a Single Station"
Authors: Bin Li, Jian Huang, Yanming Feng, Fuhong Wang, Jizhang Sang
Date Published: April 22, 2020
Date Revised: December 4, 2020
Date Accessed: September 25, 2025

Notes:

- Orbit prediction (OP) of space debris is necessary for space situational awareness (SSA), but due to the physical features of the debris, it can be hard to predict the exact trajectories of space debris
    - There is also "rapid error growth" over time that makes it hard to accurately predict the exact locations
    - A machine learning (ML)-based approach to analyzing orbital debris would be helpful to identify patterns and predict position
    - Accuracy increases with more data in the algorithm
- "More than 23 000 orbital objects larger than 10 cm in diameter tracked by the U.S. space surveillance network (SSN)"
- Other applications for ML models in the field are to:
    - Plan the optimal spacecraft position and trajectory
    - Orbit determination
    - Automatic space object characterization

Source 5: "A Machine Learning Approach to Space Debris Characterisation and Classification"
Author: James Allworth
Date Published: January 2022
Date Revised: June 30, 2022
Date Accessed: September 23, 2025

Notes:

- Problem: space debris is an increasing problem because of rise of space accessibility + difficult to remove space debris
    - Space debris has a HIGH VELOCITY
    - Current risk mitigation strategies use SSA (space situational awareness) + using orbital mechanics to PREDICT trajectories of debris
    - Info must be accurate + fast/efficient so that satellities/spacecraft should be moved
- Methodology:
    - Using neural networks with a large and well-labelled data set for increased accuracy -> apply deep learning
    - Model is transferred to a small, real-world data set to be worked upon for greater accuracy
- Results:
    - A significant step has been made towards an AUTOMATED data-focused approach to RSO (resident space object) classification from optical data
    - Machine learning techniques are effective and cost effective -> lowers barriers to entry for commercial uses in the field
    - Machine learning models that are not properly trained will not be able to handle situations/shapes/sizes of space debris data that they have not seen in training data
- Future Work:
    - Studying a classifier other than shape from light curve data (ex. size, mass, altitude, material properties)
    - Develop a network that can analyze shape and altitude -> combination of many observational techniques?
    - All future work should keep up with the current machine learning models + capabilities as they improve

Source 4: "Comparative Analysis of Resident Space Object (RSO) Detection Methods"

Authors: Vithurshan Suthakar , Aiden Alexander Sanvido, Randa Qashoa, and Regina S. K. Lee
Date Published: December 7, 2023
Date Accessed: September 20, 2025

Notes:

- Problem: rapid growth of satellites/debris in LEO - higher collision risk
- Optical images help detect debris but datasets are scarce
- Methodology:
  - Dataset - 429 low-res images collected are examined -> using classical methods, not AI
  - Comparing algorithms (AFD, MFD, PFT, Streak detection
- Results:
  - Best for on-board ships: MFD
  - Best for real-time detection: PFT
  - Best accuracy (least false negatives): Streak detection
  - Most flexible/well-rounded: MFD
- Future Work:
  - Steady image frames and stabilize them around a certain point
  - Less thresholding (diving into regions based on intensity level of pixels) to keep precision

Source 3: "ESA's Annual Space Environment Report" - DIAGRAMS I CAN USE IN HERE (page 9 - diagram of simulated predictions of LEO collisions for next 200 years)
Author:
Date Published: March 31, 2025
Date Accessed: September 19, 2025

Notes:

- More space debris in orbit than operational satellites -> poses a significant problem for near-Earth environment
- European Space Agency (ESA) publishes annual "Space Environment Report" to raise awareness about the issue of mitigating space debris in space and sustain space flight
  - Ex. requiring all ESA projects to have a post-mission lifetime limit of 25 -> 5 years

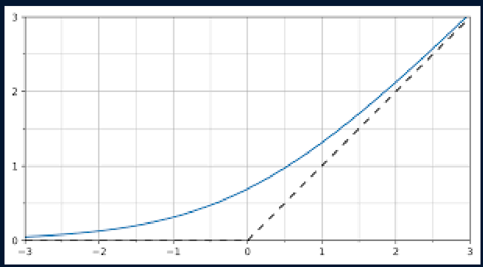
- Amount of objects, their total mass (and area) has been rising since the beginning of humankind's space age -> leads to involuntary collisions of payloads and space debris
- Improvements in space surveillance sensor capabilities in last few decades
    - Miniaturism of space systems but number of space missions have risen
- Most payloads located in 500km - 600km altitude -> this is in Low Earth Orbit

Source 2: Neural Networks Pt. 2: Backpropagation Main Ideas
Name: "Neural Networks Pt. 2: Backpropagation Main Ideas"
Author: StatQuest with Josh Starmer
Date Published: October 18, 2020
Date Accessed: September 16, 2025

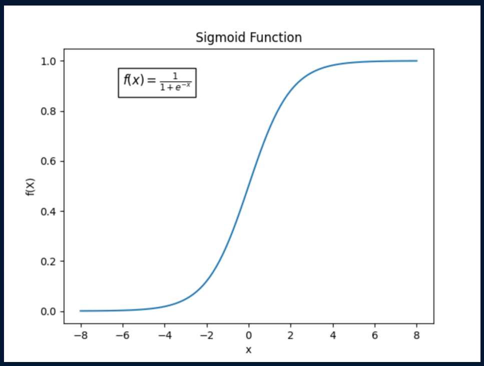
NOTES NEXT MONTH.

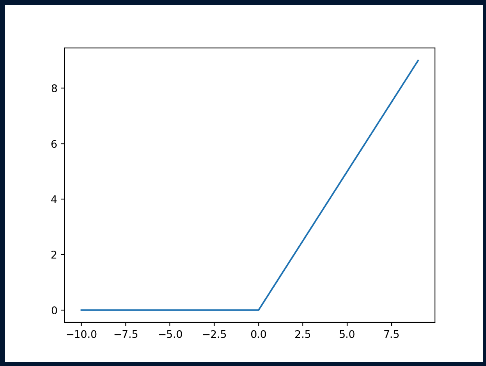
Source 1: The Essential Main Ideas of Neural Networks
Name: "The Essential Main Ideas of Neural Networks"
Author: StatQuest with Josh Starmer
Date Published: August 30, 2020
Date Accessed: September 15, 2025

Notes:
- **Neural networks**: popular algorithm in MACHINE LEARNING
    - Neural networks fit to data with BACKPROPAGATION (video part 2)
    - Many variations of neural networks since they are so complex
- Need a data set to be able to predict what the outcomes of future data will be
    - Most data can not be connected though a simple straight line -> neural networks created a **"squiggle"** through the data to find trends
- Neural networks consists of NODES and CONNECTIONS BETWEEN THE NODES (estimated parameter values from data set given -> via backpropagation)
    - Parameter values are like slope and y-int when trying to fit straight line to data -> **using backpropagation**
    - Sloped/curved lines in nodes "added together" to create uniting "squiggle"
        - Common lines:



        - CURVED: Soft Plus
        - CURVED: Sigmoid
        - BENT: Rectified Linear Unit (ReLU)





- Curved/net lines = activation functions
    - Common to use softplus/ReLU activation function
    - Connections have math operators to change data from data set into a coordinate (x-coordinate) on the activation function
        - Plug x-coordinate into math equation for the activation function to find corresponding y-value
        - Note: **log() = natural log (ln)/log base *e*** in machine learning
    - Increase/decrease input node value to get more coordinated on activation function -> plot these (y-coordinate) values into original data set & scaled y-coordinates -> makes 2 curves on og data set
        - Repeat on all hidden layer nodes to get more curves on og data set
        - Mathematical parameters (**MULTIPLICATION - WEIGHTS, ADDITION - BIASES**) to manipulate all values through neural network -> output node -> gives final answer in context of the og data (MAKES PREDICTION)
- Each neural network have **INPUT** and **OUTPUT** nodes
    - Can have 1+ input/output nodes + many connection nodes (HIDDEN LAYERS)
    - Guess how many hidden layers + nodes you need -> test -> see how well it performs -> add more hidden layers + nodes if needed
- Connections between nodes in neural network - synapses between neurons in the brain

🗓️ October Background Research

BACKGROUND RESEARCH - OCTOBER

Source 13: “Familiarization with Pandas for data science and analysis”
Authors: Akriti Thakur, Abhinav Kumar, Sachin Awasthi
Date Published: April, 2024
Date Accessed: October 21, 2025

Notes:

- Pandas is the core Python library for labelled, table-like, time-series data built on NumPy
  - Has the potential to build large files that support lots of data -> very suitable for my project since it has a lot to do with organizing large amounts of data for analysis by an ML model/algorithm
  - Pandas have indices/indexes, column names, and table operations
  - A DataFrame is a 2-D labelled table/Series is a 1-D labelled array
- Python integrates well with large data work since it has clean syntax, a large ecosystem (NumPy, SciPy, pandas, PyTorch, etc.), and lots of potentials for analysis

Source 12: “Orbit Precision Analysis of Small Man-Made Space Objects in LEO Based on Radar Tracking Measurements”
Authors: M. Kirschner, M. Weigel, R. Kahle, E. Kahr, P. Choi, K. Letsch, L. Leushacke
Date Published: 2012
Date Accessed: October 20, 2025

Notes:

- Goal of project: refine debris orbits for increased accuracy in conjunction assessments (CA), TLEs alone are not precise enough -> also used radar-based tracking data to support
  - Summarized results: after 24 hours of tracking, the radial error (the error of a satellite’s predicted distance directly away from/toward the

        Earth) was LESS THAN 40 M (more accurate than using just TLEs and radar-based data for decisions)
- More error growth with time -> small error at 1 day epoch, arc length of space object and changes in drag can increase error residuals of orbit propagation

Source 11: “Precise and Efficient Orbit Prediction in LEO with Machine Learning using Exogenous Variables”
Authors: Francisco Caldas, Cláudia Soares
Date Published: July 3, 2024
Date Accessed: October 20, 2025

Notes:
- Problem: SGP4 propagation is fast but it misses drag/perturbations -> the full numericals are slow
    - ML IS NEEDED FOR FAST AND ACCURATE ORBIT PROPAGATION
    - Project in the research paper used a feedforward neural network (FNN) -> inputs move in and produce outputs, info flows in one direction
- Idea/Project: forecast future state vectors with ML systems using past position + variables from the environment
    - Environmental factors can affect the orbit trajectories, and therefore their propagated orbits
- Prior work in the field: Hybrid SGP4 and error correction technologies with generalizations

Source 10: “Colliding Satellites: Consequences and Implications” - DIAGRAM I CAN USE IN HERE (diagram showing the spread of orbital debris over time around Earth’s orbit)
Author: David Wright
Date Published: February 26, 2009
Date Accessed: October 9, 2025

Notes:
- Iridium-33 and Cosmos-2251 collided on Feb 10, 2009, adding lots of debris into LEO space (LEO was already crowded)
    - Relative speed was around 10 km/s
    - Most debris stays concentrated near the original altitude, but over time it forms a global “shell” or ring
    - Estimated debris produced from this collision = 1 000-2 000 pieces greater than 1 cm and 60 000-120 000 pieces greater than 1 cm

- Debris 1-10 cm can destroy a satellite given the high speeds of travel and they cannot be reliably tracked
- NASA describes this band of orbit as "supercritical" -> collisions create debris faster than drag can remove it, and the population can double soon
    - Stronger rules and regulation of space-traffic management are likely needed to ensure spacecraft stay safe

Source 9: "Improving Orbit Prediction Accuracy through Supervised Machine Learning"
Authors: Hao Peng, Xiaoli Bai
Date Published: January 15, 2018
Date Revised: January 16, 2018
Date Accessed: October 1, 2025

Notes:
- Core problem: purely physics-based orbit prediction (often does not have the accuracy needed for collision avoidance)
    - Lack precision environment info + RSO properties
- Proposed solution: adding a machine learning-centered module to a physics-based pipeline that models prediction error
- Use a SVR (support vector regression) algorithm since it is robust to outliers, framework
- Validate using a simulation-based catalog (as truth data)
- Improves SSA (space situational awareness) since more observation stations and training data will help increase accuracy, benefits will stop after a certain point (must train on representative RSOs)
- The true residual error = estimated data by machine learning algorithm - true prediction error (feed this into ML algorithm to increase accuracy for future predictions)
- Time horizon examined affects the SGP4 prediction growth -> increased time leads to increased error growth

Source 8: "Understanding long-term orbital debris population dynamics"
Author: Hugh G. Lewis
Date Published: September 1, 2020
Date Accessed: September 15, 2025

Notes:

- Originally linear growth of debris population in Earth's orbit -> has been increasing EXPONENTIALLY
    - This growth is mainly due to the build up of objects/debris at 1200-1500 km above Earth, even with few launches of spacecraft to that band
    - Below 700 km there is lots of collision activity with many orbits crossing leading to collisions
    - Exponential increase in orbital debris is expected to stay in the long-term
    - Over the next 1000 years, growth of debris/objects in Earth's orbit is predicted to roughly double
    - Predicted interval between catastrophic collisions expected to drop from 49 years to 6.6 years over the next 1000 years
- Ultra-long/long-run predictions are necessary to reveal the EXPONENTIAL GROWTH TRENDS OF DEBRIS POPULATIONS

Source 7: "Machine learning from theory to algorithms: An overview"
Authors: Jafar Alzubi, Anand Nayyar, Akshi Kumar
Date Published: November 1, 2018
Date Accessed: September 15, 2025

Notes:
- The aim of machine learning algorithms is to adapt to the data given and create patterns/connections between data
    - The success of an algorithm is measured by improving accuracy through EXPERIENCE
    - Big Data is accelerating the adoption of ML systems/algorithms
- ML is multidisciplinary -> tied to computational statistics and mathematical optimization
    - Some other applications of ML systems (spam, fraud, medical, traffic, trading, recommendations)
- Supervised learning = learning from labeled samples/examples
- Unsupervised learning = finding hidden structure within unlabelled data
- Reinforcement learning = trial-and-error with a critic (feedback on correct/incorrect)
- Ensemble learning = combining many different types of models to reduce bias and variance (ex. Random Forest method -> uses Decision Trees)

Other Research (Not Research Papers):

Website Links:
https://www.geeksforgeeks.org/machine-learning/supervised-vs-reinforcement-vs-unsupervised/

https://www.mathworks.com/help/aerotbx/ug/orbit-pop-algorithms.html

Google Searches for general info -> Website Links:

- Supervised learning: uses labeled data to teach an algorithm to make predictions by learning a mapping from inputs to **known outputs**
    - Comparable to a teacher
- Reinforcement learning: teaches a model to make decisions in an environment through **trial and error**, learning by maximizing cumulative rewards rather than using explicit labels
- Propagating an orbit: calculating an object's future position and velocity over time based on its initial conditions and the forces acting upon it

🗓️ November Background Research

BACKGROUND RESEARCH - NOVEMBER

Source 15: "Improved Orbit Propagator Integrated with SGP4 and Machine Learning”
Authors: Zac Yung-Chun Liu, Scott Tarlow, Mohammad Akbar, et al.
Date Published: November 17, 2025
Date Accessed: November 2025

Notes:

- SGP4’s limitations: widely used for fast satellite orbit predictions but this physics-based propagation system is only accurate for short-term satellite tracking (around 12 hours in advance)
- The authors created a hybrid SGP4 and machine-learning approach in the hopes of improving orbit predictions
  - Did not replace SGP4, instead ran it to get outputs then put a ML system over the outputs to reduce the error behaviour
  - Used 2 types of ML models (autoencoder neural network, random forest)
- SGP4 errors come from simplifying assumptions -> does not account for ALL forces acting on an object in space (like atmospheric drag changes, etc.)
  - Over days and weeks, these errors lead to error growth in propagated states
  - ML model was trained to capture the “time-series pattern” of these errors
- Needed 3 years of past TLE data records for each satellite they used as a training dataset -> this data can be organized into a Pandas DataFrame in

Python (can align by timestamps and compute SGP4 predictions versus the actual position of the satellite)

- Results: tested on 3 different satellites, the hybrid SGP4 + ML approach greatly improved orbit prediction accuracy
  - About a 20%-30% reduction in prediction error over a 30-day period compared to JUST USING SGP4 ALONE
- *****Python implementation:** this approach/methodology is implementable in code (using a Python program to integrate a SGP4 propagation system -> using a library like ***python-sgp4***, Pandas for data handling, and TensorFlow for a machine-learning neural network)
  - The Python system is packageable, and Python was used to "glue together" all the different pieces of this project
    - Reading TLE files -> Pandas DataFrame -> running SGP4 in Python -> training model (TensorFlow) -> making new, more accurate predictions

Source 14: "Space Weather Environment During the SpaceX Starlink Satellite Loss in February 2022"
Authors: Tzu-Wei Fang, Adam Kubaryk, David Goldstein, et al.
Date Published: October 24, 2022
Date Accessed: November 7, 2025

Notes:

- Peer-reviewed research paper that analyzed the loss of 38/49 Starlink satellites on Feb. 3, 2022
  - Failures mainly because of upper-atmosphere density during a geomagnetic storm
- Starlink team tried to recreate the density environment of satellite orbits using SIMULATIONS (used observations, predictions/forecasts, and physics-based calculations)
- "Minor to moderate" geomagnetic storm was strong enough to produce 50-125% increase in density at altitudes of 200-400 km (INCREASE DRAG ON SATELLITES)
  - The Kp index (3-hour geomagnetic activity index) reached Kp=5 (stayed at this level for 6 hours on Feb. 3 and Feb. 4) and F10.7 values were higher than normal

    - This is classified as a "minor geomagnetic storm" on the National Oceanic and Atmospheric Administration (NOAA) scales
    - Even though it was only a minor geomagnetic storm, the long **duration** (6 hours two days in a row) was significant to satellites in LEO and VLEO -> this is where the Starlink satellites were
- During the event, DRAG was recorded to be much higher than expected -> only 11/49 satellites could be reoriented to a "safe" position and altitude
- F10.7 (the 10.7cm solar radio flux) is an index for representing solar extreme UV-output -> higher F10.7 values means more UV energy and a stronger heating of the upper atmosphere
    - This affects orbits!
    - Upper-atmosphere expansion events (driven by space weather changes) are critical to measure and account for since there are many large satellite constellations in LEO/VLEO - can also lead to drag uncertainty

Other Research (Not Research Papers):

Website Link:
https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.html?utm_source=chatgpt.com
Accessed: November 17, 2025

Notes

- A Pandas DataFrame is a 2-D, size mutable table with LABELLED rows and columns
    - Table (DataFrame) can hold mixed data types
    - DataFrames behave like a dictionary of aligned SERIES objects
        - Can be constructed from NumPy arrays, other DataFrames, **dicts of arrays series**
    - ***Index*** parameters lets you set column labels and can be used to define names
    - Lots of data cleaning and sorting tools are built into Pandas in Python and can be applied to the DataFrame and the data inside it

    - ***.to_numpy()*** or ***.values*** parameters convert a DataFrame into raw arrays that can be FED DIRECTLY INTO MACHINE-LEARNING LIBRARIES -> TENSORFLOW, PYTORCH, ETC.

Website Link:
https://pypi.org/project/sgp4/?utm_source=chatgpt.com
Accessed: November 17, 2025

Notes

- The SGP4 Python package is used to track Earth satellites given TLE data (or OMM data - I will not use this for my project)
- Compiles the official C++ SGP4 code from famous "Revisiting Spacetrack Report #3) -> Vallado
- Position outputs are given as x, y, z coordinates in km in the Earth-centered time frame - NOT AS LAT/LONG
    - Another astronomy library (ex. Skyfield) can be used to convert the output coordinates of the SGP4 propagations into other frames of reference (J2000, ITRF, lat/long)
- The package provides helpers (like ***jday***) to convert calendar dates into JULIAN DATES

Website Link:
https://spaceweather.gc.ca/forecast-prevision/solar-solaire/solarflux/sx-3-en.php
Accessed: November 7, 2025

Notes

- The Natural Resources Canada/Space Weather Canada website has 2 databases of solar flux data - 10.7cm Flux measurements , and daily record of the flux outputs
    - 10.7cm Flux measurements are expressed w/ 3 values: observed, adjusted, URSI Series D values
    - Observed value: number measured by the solar radio telescope -> affected by the level of solar activity and the variable distance between the Earth and the Sun
    - Adjusted value: calculation performed on the data to correct for variations in the Earth-Sun distance (uses average distance)
    - Series D Flux: a scaling factor of 0.9 given to the data so it can be fitted onto a spectrum -> not usually used anymore

    - 3 flux determinations are made PER DAY, quantities are separated by commas
    - Measurements in the 10.7cm Solar Flux data: Julian Day of the measurement, Carrington Rotation Number, year, month, day, observed flux, adjusted flux, Series D flux

Website Link:
https://www.swpc.noaa.gov/products/planetary-k-index
Accessed: November 7, 2025

Notes
- The K-index (Planetary K-index) characterizes the magnitude of geomagnetic storms
    - Indicator of DISTURBANCES in Earth's magnetic field, used to determine whether geomagnetic alerts/warnings need to be issued
    - People/things affected by geomagnetic storms: electrical power grid, spacecraft, spacecraft operators, aurora observers, users of radio signals that go through ionosphere
- Measures the disturbances in the horizontal component of the Earth's magnetic field (uses integers 0-9, 1 is calm and 5+ is a geomagnetic storm)
- Observed using a magnetometer, and the maximum fluctuation data is recorded in 3-hour intervals or one line per day (depending on the data source)

🗓️ December Background Research

BACKGROUND RESEARCH - DECEMBER

Source 17: "Analysis of the Application Efficiency of TensorFlow and PyTorch in Convolutional Neural Network"
Authors: Ovidiu-Constantin Novac, Mihai Cristian Chirodea, Cornelia Mihaela Novac, et al.
Date Published: November 16, 2022
Date Accessed: Dec. 11, 2025
https://pmc.ncbi.nlm.nih.gov/articles/PMC9699128/#:~:text=users,the%20field%20of%20artificial%20intelligence

Notes:

- TensorFlow and PyTorch are both popular libraries for developing deep learning models + both are widely used in the AI field
- Differences in design:
    - TensorFlow: uses **computation graphs** (users can define the whole neural network first) -> more user-friendly and a more mature ecosystem
    - PyTorch: uses **dynamic graphs** (graphs can be modified on the fly) -> more flexibility but a smaller activity community, models train FASTER
    - Ex. 16hr 59 mins in PyTorch and 21hr 57 mins in TensorFlow on the same hardware
    - Ex. Very similar accuracy rates of both models trained on the same hardware -> framework used does not significantly change accuracy
- Both have similar capabilities (predefined layers, optimizers, etc.)
- Both have been used in satellite and space-related ML tasks + both can integrate well with space-related Python libraries

Source 16: "TensorFlow: A system for large-scale machine learning”
Authors: Mart´ın Abadi, Paul Barham, Jianmin Chen, et al.
Date Published: November 2–4, 2016
Date Accessed: Dec. 5, 2025

Notes:

- TensorFlow: open-source framework for ML -> developed by Google
  - Designed to handle large-scale computation across different devices
- Represents computations as **dataflow graphs**
  - “Nodes” are operations (like math functions) and edges are data arrays called *tensors* flowing between them
- Built to **train deep neural networks** on large datasets -> then run those trained models for predictions efficiently
- Widely used in both industry and research
- Typical neural network model = defined as a directed graph of layers ending in a *loss function* (which measures prediction error)
- Each layer of the neural network performs a set of mathematical operations
- TensorFlow supports many types of neural networks -> very flexible for various ML tasks
- Commonly used for **supervised learning** -> you can train a neural network on labeled data
  - TensorFlow will then adjust the model’s weights to **minimize the loss (error)** over the training examples

🗓️ January Background Research

BACKGROUND RESEARCH - JANUARY

Source 18: "Improved Orbital Propagator Integrated with SGP4 and Machine Learning"
Authors: Zac Yung-Chun Liu, Scott Tarlow, Mohammad Akbar, et al.
Date Published: July 9, 2021
Date Accessed: Jan. 24, 2026

Notes:

- SGP4: a standard physics-based algorithm used to propagate (predict) satellite orbits
    - Its accuracy diminishes for long-term predictions (beyond about 12 hours) due to unmodeled forces and approximations
- Researchers developed a **hybrid model** that combines the traditional SGP4 orbit propagator with ML techniques to correct SGP4's errors
    - Research noticed that SGP4 residual errors over time have patterns (like missing physics like drag or gravitation cause systematic deviations)
        - ML part is trained to recognize error patterns from historical data training sets
    - Used a random forest regressor (predicted error corrections for SGP4) and an autoencoder
    - Supervised learning approach training on historical orbit data where "inputs" were SGP4's initial predictions, and "outputs" were the real positions of the space objects/residual error -> ML model learned how much SGP4 predictions differ from actual positions
- Results: hybrid ML-SGP4 model significantly improved orbit prediction accuracy compared to SGP4 alone
    - Over a 30-day prediction period, the error was reduced by around 20% to 20% with ML corrections (ML model predicts residual error offset and applies it to make the output closer to the truth position)

- Study used 3 years of TLE data for each satellite to train the ML model on for an increased accuracy
- ML corrected SGP4 values generate about as fast as normal SGP4 values, making this a useful tool for operational use in companies + mission planning software

📆 February Background Research

BACKGROUND RESEARCH - FEBRUARY

Other Research (Not Research Papers):

Website Link:
https://keras.io/api/losses/
Accessed: February 7, 2026

Notes

- Loss function: computers a single number that measures how wrong/inaccurate your model's predictions are
  - ML training tries to **minimize this number** by adjusting model's weight
  - In Keras/TensorFlow, the loss is the key thing you set in `model.compile(loss=...)`
- Loss is computed from **`y_true`** **(ground truth)** and **`y_pred`** **(model output)**
  - For my project, `y_true` is the **true residual correction** [dx,dy,dz], and `y_pred` is the model's predicted correction vector
- Loss is usually computed **per sample**, then combined ("reduced") across the batch (usually averaged)
  - Common reduction is **mean over batch** ("sum_over_batch_size"), so the loss doesn't depend on batch size
- Different losses behave differently with **outliers** (rare huge errors)
  - **MSE (Mean Squared Error):** squares errors, can get yanked around by outliers -> *least stable*
  - **MAE (Mean Absolute Error):** uses absolute value → more robust to outliers, gradients are "flatter"

    - **Huber loss:** hybrid (acts like MSE for small errors, like MAE for big ones), best used for mostly normal data with a few outliers
- Loss is what training optimizes, should also track **metrics** I care about in the project (like 3D error magnitude)

🗓️ March Background Research

BACKGROUND RESEARCH - MARCH

https://science.nasa.gov/learn/basics-of-space-flight/chapter2-2/

📆 April Background Research

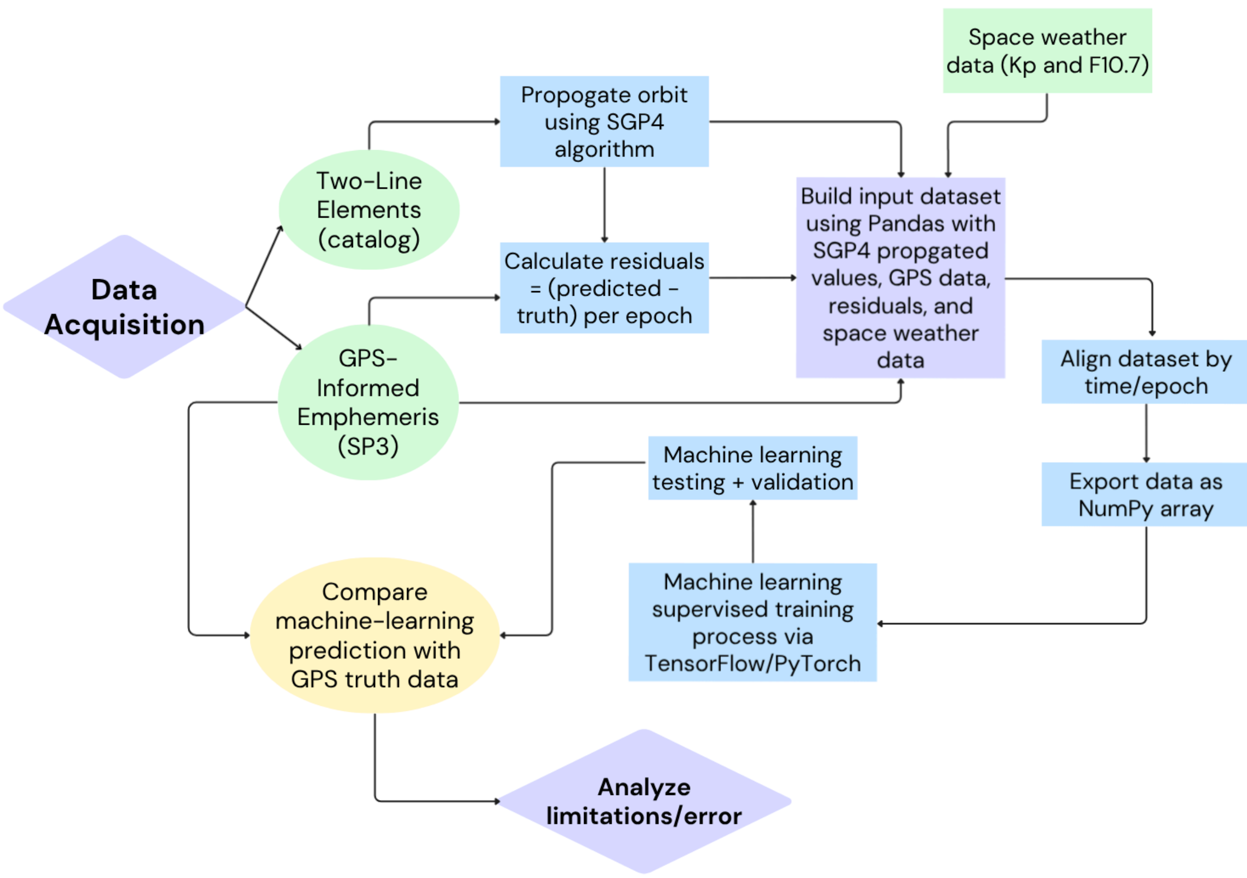
BACKGROUND RESEARCH - APRIL

Methodology

METHODOLOGY

Helpful Article Describing Break-Down of Tentative Methodology Steps:
https://mads-hatters.github.io/#:~:text=Data%20Source%201%20Space,Line%20Element%20%28TLE

Finalized Methodology Flowchart in Research Proposal:



Finalized Methodology Flowchart on my Science Fair poster (OP #2 presentation poster):

Methodology

1) Data Collection

- 26 weekly "TLE packets" of CASSIOPE satellite
- CASSIOPE GPS truth data
- Space-weather (Kp, F10.7) aligned by time

2) Baseline Propagation

- Use SGP4 to propagate each TLE packet forward 2 weeks
- Convert to ITRF at each epoch

3) Computing Residual Error

- Match SGP4 baseline epochs to GPS truth data epochs
- Residual error = |truth-SGP4| (km)

4) Build ML Dataset

- Align features by same time/epoch
- Features: Baseline states (SGP4), space weather, time since epoch

5) Train ML Model on Residuals

- Use Neural Network ML to predict residual components given input data features
- Training using 20 TLE packets

6) Apply Time-Gated Correction

- corrected = SGP4 – a(t) * c_pred
- a(t) = 0 until day 7
- a(t) = 1 after day 7
- Goal: stabilize early and correct later

7) ML Validation + Results

- Validate ML model on 6 unseen TLE packets
- Compare baseline SGP4 propagation error to after-ML error

📄 Major Assessments

MAJOR ASSESSMENTS

See individual subtabs for the major assessments.

LINKS:
Research proposal: Applied Science Project Research Proposal
Oral presentation: ASP Research Proposal Oral Presentation - Elise Protti

🛰️ Research Proposal

MACHINE-LEARNING ENHANCED ORBIT PROPAGATION: IMPROVING LOW EARTH ORBIT PREDICTION USING TLE AND GPS DATA

by

Elise Protti

Andrew Howarth, University of Calgary Department of Physics and Astronomy

Research Proposal

Applied Science Project

October 22, 2025

Introduction

The surge in the number of pieces of orbital debris in Earth's orbit has become an increasing problem ever since humanity started exploring and expanding into space at the start of the space age in 1957 [1], [2], sending up commercial operators and large constellations of satellites to provide data about factors such as climate, atmospheric conditions, and navigation on Earth [3]. The accumulation of a large population of debris surrounding the Earth is making it extremely difficult to maintain sustainability in space in the long-run due to the dangers of space debris striking satellites and spacecraft. It is estimated that there are around 34,000 trackable objects that are 10 cm or larger in Earth's orbit, posing significant challenges to space travel and development in space [2]. To illustrate the scale as well as composition of the orbital environment targeted by the project, the European Space Agency's 2025 annual report includes a useful figure showing statistics of growth of cataloged objects across all orbital classes over time. These objects are differentiated by object type– such as unidentified objects, rocket mission related objects, rocket debris, rocket fragmentation debris, rocky bodies, payload mission related objects, payload debris, payload fragmentation debris, and payloads – and the trend of the data illuminates that after decades of gradual increase in the number of objects in orbit, the curve steepens dramatically in the last decade as large constellations of satellites and major collisions add active spacecraft and debris, respectively.

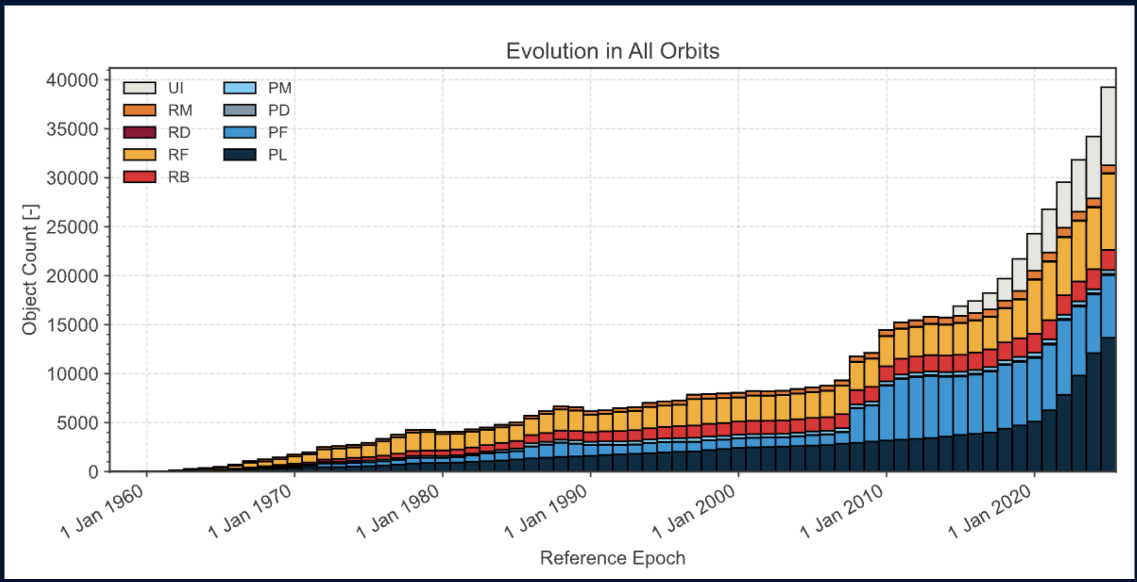


Fig. 1. Evolution of object counts in all of Earth's orbits stacked by object class from 1960 to present. The objects are differentiated by object class – unidentified objects, rocket mission related objects, rocket debris, rocket fragmentation debris, rocky bodies, payload mission related objects, payload debris, payload fragmentation debris, and payloads, respectively – and show a gradual then steep increase in the number of objects in Earth's orbit in the last decade. Adapted from [3].

Figure 1 underscores why automated, machine learning-assisted predictions and orbit propagations are increasingly necessary for timely and accurate risk assessment in space. Due to orbital debris' inherent lack of navigational systems, in contrast to satellites and spacecraft, uncertainty arises about the exact trajectories of these objects, adding another level of risk to human-made spacecraft. Several incidents have happened in history where satellites and spacecraft have collided with objects in space, in turn creating a larger, vast field of debris which keeps the cycle of debris in Earth's orbit and increases the risk level even more [4]. An example of this was the February 10, 2009 collision of the U.S. Iridium 33 communication satellite with the derelict Russian Cosmos 2251 communication satellite both orbiting the Earth at an altitude of around 790 km [5], [6]. Their collision speed was about 11 km/s, and this event resulted in the creation of two new major debris fields that contributed to the dispersal of orbital debris around Low Earth Orbit (LEO) space [5], [6].

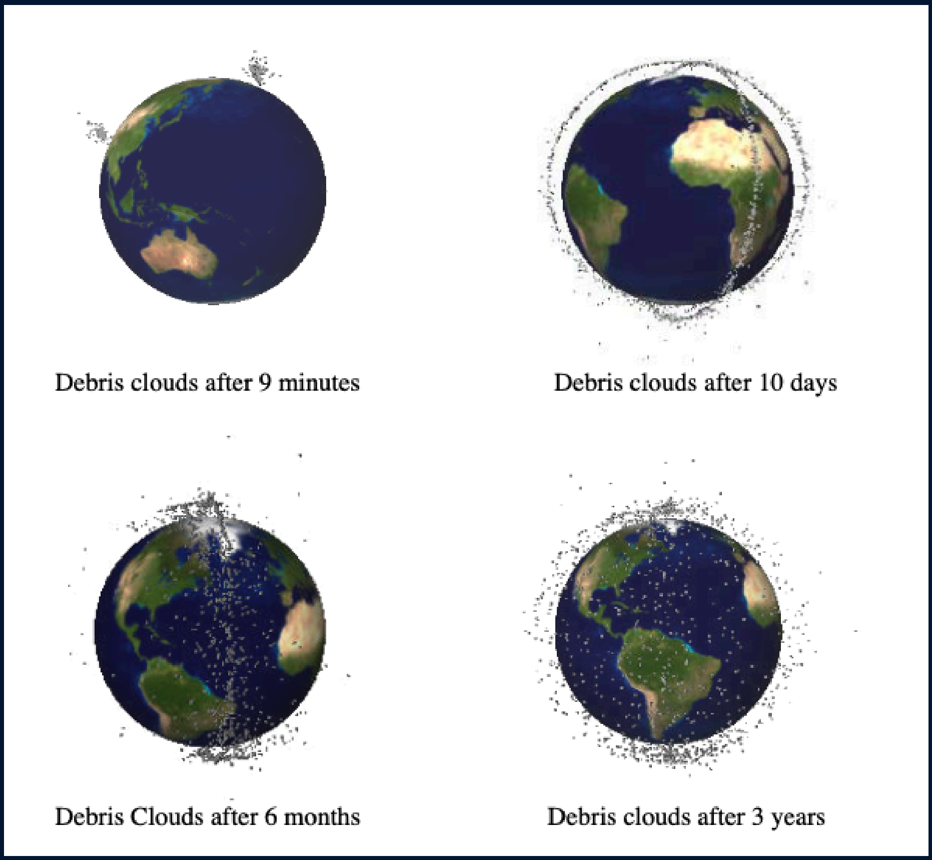


Fig. 2. Timestamped projections diagram showing the general pattern of the spread of debris clouds and their distribution approximately 9 minutes, 10 days, 6 months, and 3 years after a major collision resulting in debris formation. Adapted from [5].

Figure 2 aids in visually showing how space debris clouds resulting from catastrophic collisions in LEO can spread and diffuse quickly, illuminating how the large amount of debris in orbit is a sizable risk to human spacecraft and satellites. Therefore, the need for accurate and prompt orbit propagation systems to ensure all spacecraft stay intact is more crucial than ever in the modern age.

Having accurate predictions of satellite trajectories through the use of detection algorithms is a crucial part of space-situational awareness (SSA), which refers to the knowledge of specific locations of objects and hazards in space, and it ensures spacecraft and satellites are kept safe while in orbit [7]. Recent improvements in the sensor capabilities of space surveillance technologies have even made it possible to track and catalog smaller pieces of debris more than ever before [3], which opens up the opportunity for the orbits of these pieces of debris to be analyzed and predicted using orbit propagation techniques.

The current popular method of propagating and predicting orbital trajectories involves a physics-based method, which can produce volatile predictions due to the many factors of space, such as atmospheric drag, solar radiation pressure, and Earth's gravity [4]. The surveillance equipment required to collect these precise measurements is very expensive and the process of data acquisition is time consuming, making it almost impossible to have timely orbital predictions [4]. Additionally, the errors from this type of orbital prediction – which typically uses Simplified General Perturbations-4 (SGP4) propagator systems – increase very rapidly as time goes on since over-simplifications are made by the system, especially when trying to predict the orbit of an object in LEO over extended periods of time or long distances [8], [9]. Therefore, physics-based orbital propagation is a pressing issue in the modern-day fields of aerospace development and satellite tracking since there is a very large potential for error, and a few meters in LEO space could make all the difference between a close-call encounter between pieces of debris or a full-on collision [8].

Machine-learning methods offer a promising new approach to addressing this issue, as the predictions of the orbits can be made without the explicit data points of atmospheric drag, gravitational forces, and more [4]. In contrast to physics-based orbital propagation methods, machine-learning algorithms work similar to a human brain's neural networks by taking in large amounts of data, recognizing patterns, and using past data to predict future events based on these patterns [4]. Specifically, supervised learning is a machine learning training technique where labelled input data with corresponding known outputs is inserted into a machine learning algorithm [4], [9]. This method of training is very effective since the algorithm is able to devise patterns and relationships within the data to propagate future orbits [4], [9].

A key type of data that is crucial in this machine-learning process are Two-Line Element (TLE) sets. TLEs are an example of a public, open-source data record for resident space objects (RSOs), encoding an object's orbit at a specific epoch (moment in time), and can also propagated with SGP4 to predict their future position [8], [10]. TLEs exist for thousands of cataloged objects, always appearing in a consistent format, making them an extremely useful, standardized type of data that can be sorted through easily by a machine-learning system [4]. Leveraging these consistent, widely available records of an object's position, paired with their SGP4 propagated states will offer excellent input data for a machine-learning algorithm.

Prior work in this field includes scientists using supervised and reinforced machine learning algorithms to introduce a machine-learning-centered strategy into traditional physics-based propagation methods [4]. Most of these systems utilize known "truth" data which is compared with ground-based data to make predictions about the level of error of the physics-based SGP4-propagated data [4]. Data types historically used for this step have included simulation-based space catalog environments, RSO radar-based observational data, image-based data, and more [4], [10]. The accuracy and usefulness of these machine-learning systems have depended on many factors which set each research study apart, such as: the type of machine-learning algorithm used, the type of machine-learning style used, the program or development environment used to implement it, the accuracy of the data inputted, and the time range examined [4], [8], [10].

Different machine-learning systems use different classifiers to approach data sets, such as random decision forests, which combine multiple decision "trees" to improve the accuracy and robustness of prediction [8], and support vector machines, which plot points in n-dimensional space based on the number of features in the dataset to come to conclusions [10], [11]. The differences in these methods can translate to the accuracy of the results given. In general, though, the prediction error of orbital trajectories has decreased due to the general usefulness and accuracy of machine-learned augmented predictions that help minimize the physics-based prediction errors [4]. Since the machine-learning field and information about the increasing amounts of orbital debris in LEO space is relatively new, many scientists hint that more research needs to be done and more machine-learning algorithms need to be tested for indisputable conclusions to be made on the subject area [4].

Therefore, this innovative study is aimed at specifically targeting the error growth of physics-based orbit propagation methods. The study will focus on training a supervised machine-learning model (with a neural-network baseline) using satellites that provide precise "truth" orbits via GPS data in comparison to their predicted future positions propagated by a standard SGP4 algorithm. By the machine-learning algorithm systematically learning the residual error between the satellites' predicted and observed states, the model can possibly increase the accuracy of orbit propagation by relying on patterns in the data to augment physics-based orbit propagation methods, which proves to be inaccurate over long periods of time and distance [12]. This same algorithm may also be applied to debris objects that are forced to rely only on cataloged data, as they do not have the ability to have their exact position "truths" determined. However, it is important to note that data gathered from radar-based techniques on ground observatories will never be as accurate as GPS data. GPS data is generally accurate to decimeter or centimeter levels, whereas radar-based data is only accurate to a scale of meters because of the nature of the data collection process [13].

Research Questions

Primary Research Question:

- The primary research question that the study is addressing: can a supervised machine-learning correction model, trained on residuals from SGP4 propagated orbits (from inputted TLE data) and basic space weather information, significantly reduce multi-day or extended orbit prediction error for LEO objects?
- Furthermore, does this improvement translate to a noticeable positive difference in accuracy and timeliness of corresponding debris conjunction assessments (CAs)?

Secondary Research Questions:

- Does the type of space object examined by the machine-learning algorithm – for example, the CASSIOPE or Swarm A/B/C satellites – affect the accuracy of the algorithm given the differences in orbits, or will the algorithm output a similar accuracy regardless of the actual values of the data?
- Do AI-corrected predictions reduce the number of “false alarms” or decrease the volatility of near-miss distances in LEO space?

Objectives

General Objective:

The main, general objective of this study is to determine whether a supervised machine-learning correction model can reduce multi-day LEO orbit propagation prediction error compared to the standard physics-based SGP4 propagation method. If successful, this would improve the accuracy, usefulness, and timeliness of CAs of orbital debris.

Short-Term Objectives:

1. Acquire, compile, and align a dataset of historical catalog records and GPS “truth” ephemerides for the satellite of interest (CASSIOPE and Swarm A/B/C satellites).
2. Build a baseline for the data. Run SGP4 orbit propagation, measuring the average position of the satellite over different periods of time, and compute the residuals between the SGP4 orbital predictions and the GPS “truth” data.
3. Engineer an input data set with categories of data (such as recent propagated states, standardized time, GPS truth data, residual calculations, and basic space weather elements) to input into the supervised machine-learning model using Pandas in Python.
4. Conduct fair tests with the machine-learning model, evaluating outputs on time periods not given by the initial ephemerides, and compare the errors to the SGP4 baseline as well as GPS truth to analyze accuracy. Apply the machine-learning algorithm to CASSIOPE and Swarm A/B/C satellites since they have varied orbits.
5. Report results of the algorithm tests and their resulting accuracies, along with any limitations found in the process, in a research paper at the end of the school year.

Long-Term Objectives:

1. Apply the machine-learning algorithm to orbital debris instead of just satellites. Using the trained neural network-based model that was trained on the SGP4-propagated data from satellites and checked with their GPS truth, the project will strive to apply this algorithm to debris objects that only have cataloged TLE data, and measure the accuracy of the SGP4 propagated orbits versus the machine-learning predicted ones.
2. Broaden the coverage of the project by applying the algorithm on more satellites and space objects with different orbit shapes and trajectories, to test whether the model is still effective when the objects analyzed are new or different, therefore making sure the model is not overfitted to one specific space object.

Methodology

Data Acquisition:

The first step of the project and the first short-term objective is to gather the raw data that will be inputted into the machine-learning algorithm. The two types of data that need to be gathered are TLEs and high-precision GPS truth data (from onboard navigational systems), and they will be post-processed into ephemerides for the specific satellites of study. For the CASSIOPE and the Swarm A/B/C satellites, a continuous span of historical data and the matching GPS ephemerides for the same times is necessary so that residuals (error) between them can be calculated later on in the project [7]. While the cataloged data is in the form of TLEs, GPS satellite orbit data for the satellites is organized into the standardized Standard Product #3 (SP3) format which provides position and velocity data at a given time [14].

After the data for the specific time ranges is gathered, the data will be organized into a structured spreadsheet or a Comma-Separated Values (CSV) file, which is a plain-text table that can be integrated into code like Python (and integrated using Pandas) so it can be easily manipulated and also understood by a machine-learning algorithm [15]. The files will have sections for names, date ranges, data sources, and notes so all the data stays consistent and organized [15].

Data Preparation and Residuals Calculations:

The next step is to compare the predicted position of the satellite versus its true position at the exact same time/epoch. For each piece of TLE data, SPG4 will be used to propagate the satellite's position at the timestamps present in the GPS precise position ephemeris [10]. These two data streams will then be time-aligned so that each epoch will have an SGP4-predicted position and a GPS true position. Having these two types of data then makes it possible to determine the position error/residuals in the corresponding epochs [16]. The process of calculating the residuals entails finding the difference between the spacecraft's position described by the SGP4 prediction data and the GPS measurements at a certain epoch.

There are many methods used for calculating residuals in a dataset in the astrophysics field, but two popular methods are: using a high-level Python program [15], [16], or using NASA's General Mission Analysis Tool (GMAT) [17], [18]. A Python program can be leveraged to post-process orbital data from different data types and compute the difference between them, utilizing Python tools such as Astropy or NumPy [15], [16]. Alternatively, a software package, such as NASA's GMAT, can do these calculations automatedly and efficiently [17], [18].

Building Input Data Set:

The input data set will be built into a single table (or Python DataFrame) likely using Pandas, a Python library tool for data manipulation [15]. Pandas treats data similarly to working with a spreadsheet with labeled columns and rows, and it also integrates easily into code [12], [15], [19]. Each row in the table of the Pandas DataFrame will represent one comparison between the data types (the matching SGP4 predicted position and "truth" data at the same time or epoch), and can even be sorted by time using Pandas for easy comparison [15]. Other information such as residuals and simple space weather measurements can be stored in the Pandas DataFrame [15], [19].

Space weather measurements affect the orbit or space objects, so including some in machine-learning training will hopefully increase accuracy. Some parameters relating to magnetic activity level (using an index called Kp) and past/predicted solar activity level (using an index called F10.7) can be added to the input data set to potentially increase the accuracy of the orbit propagation by the machine-learning algorithm [4], [9].

In short, the basic, logistical steps of building the input data set are: load the input data into Pandas DataFrame (which makes the data easy to sort and manipulate) [15], then export this table of information as a NumPy array that a machine-learning algorithm can read directly [19]. The importance of using Pandas in conjunction with Python is that Pandas can pipeline the raw, sorted information to the machine-learning algorithm for use in the training process [15].

Machine-Learning Model Training:

The machine-learning tasks in the project will be to predict position corrections to augment the physics-based SGP4 propagation at a chosen time, rather than predict every orbital dynamics factor relating to a spacecraft or piece of debris, which will theoretically increase efficiency and precision of the predictions [4], [20]. For the machine-learning training section of the methodology, the training will be done in Python, using a machine-learning library like TensorFlow or PyTorch [9], [21]. Since Pandas is compatible with Python and allows for standardized data to be arranged in a simple manner, the acquired and calculated data – once converted into arrays – can be inputted into TensorFlow [21] or PyTorch's [9] machine-learning model training function. The machine-learning model will act as a small neural network that predicts the orbit correction at different time intervals. It works by taking input arrays from Pandas, learning the mapping from inputs to corrections, then outputting its predictions [21]. It is crucial that the algorithm is also checked using data not given in the original dataset so that the test can be fair and ensure that the model is not just memorizing the training ephemeris [9].

Machine-Learning Validation and Further Testing:

After the machine-learning training is complete, the model will then be evaluated on future periods that were not part of the initial dataset [9]. This means that training for the model will occur using past, historical data; however, the validation and testing of the model will use data from later months. This is crucial so that the model's behaviour with new data can be examined [9].

The baseline physics-based orbital propagation generated by the SGP4 system will then be checked with the machine-learning augmented position-corrected prediction, to determine the percent accuracy of the model in comparison to the true GPS position of the satellite.

For further testing, the study will examine how the machine-learning algorithm performs when it is tested on a satellite that is left out of the training set (for example, only training on the CASSIOPE and Swarm A satellite, but then testing the algorithm on the Swarm B satellite). This process will hopefully give conclusive findings about whether the methods of the machine-learning algorithm can still be applicable to objects of different orbits. Eventually, the model can be applied to other objects in space to determine their corrected orbit trajectories without knowing their exact truth data. If the model performs well on making accurate augmented predictions of trajectories, then real-world problems concerning orbital debris collisions and near-misses can be improved, since many spacecraft and debris lack GPS truth data.

Analyze Limitations and Outliers After Comparison with Truth Data:

In the case that errors spike in the results/outputs of the machine-learning algorithm, the data collected, environmental causes, and model-related causes should be examined.

In the collected data, there could be discrepancies/errors resulting from large time gaps in catalog updates for the TLE data, gaps in the GPS truth ephemeris, or time frames being mismatched in the Pandas DataFrame. Errors spotted in this data collection process can be remedied by logically checking if there are gaps in the data, and by checking that all prediction and truth data that have the same timestamp are matched up correctly, which can be a challenge due to the differences in their formats.

Environmental causes such as changes in solar activity or atmospheric drag can greatly alter the data, leading to changes in trajectory that may not be expected by the machine-learning algorithm [4], [9]. Inputting normal versus disturbed environmental space conditions into the machine-learning algorithm can likely aid in pinpointing the exact cause of the anomalies in the corrected-prediction output [11], and give an indication as to whether the model struggles with these changes in space conditions.

Lastly, model-related errors may occur if the machine-learning model learns patterns that are too specific to one space object, then applies the wrong patterns to other objects that it is examining in a process called “overfitting” [8], [20]. The crucial step of leaving some satellites/spacecraft out of the initial dataset is meant to catch this error so it can be identified before being applied to objects that do not have GPS truth data (like orbital debris).

Timestamped issues can also be created in the project logbook throughout this process of searching for limitations, outliers, and anomalies in the data so that the suspected cause and object of interest can be recorded for analysis. Additionally, predictions with a low accuracy that drop below a certain consistency threshold (which will be determined during the machine-learning validation process once more context is known about the accuracies of the predictions) can be flagged by the algorithm as needing to be analyzed for error.

Methodology Flowchart:

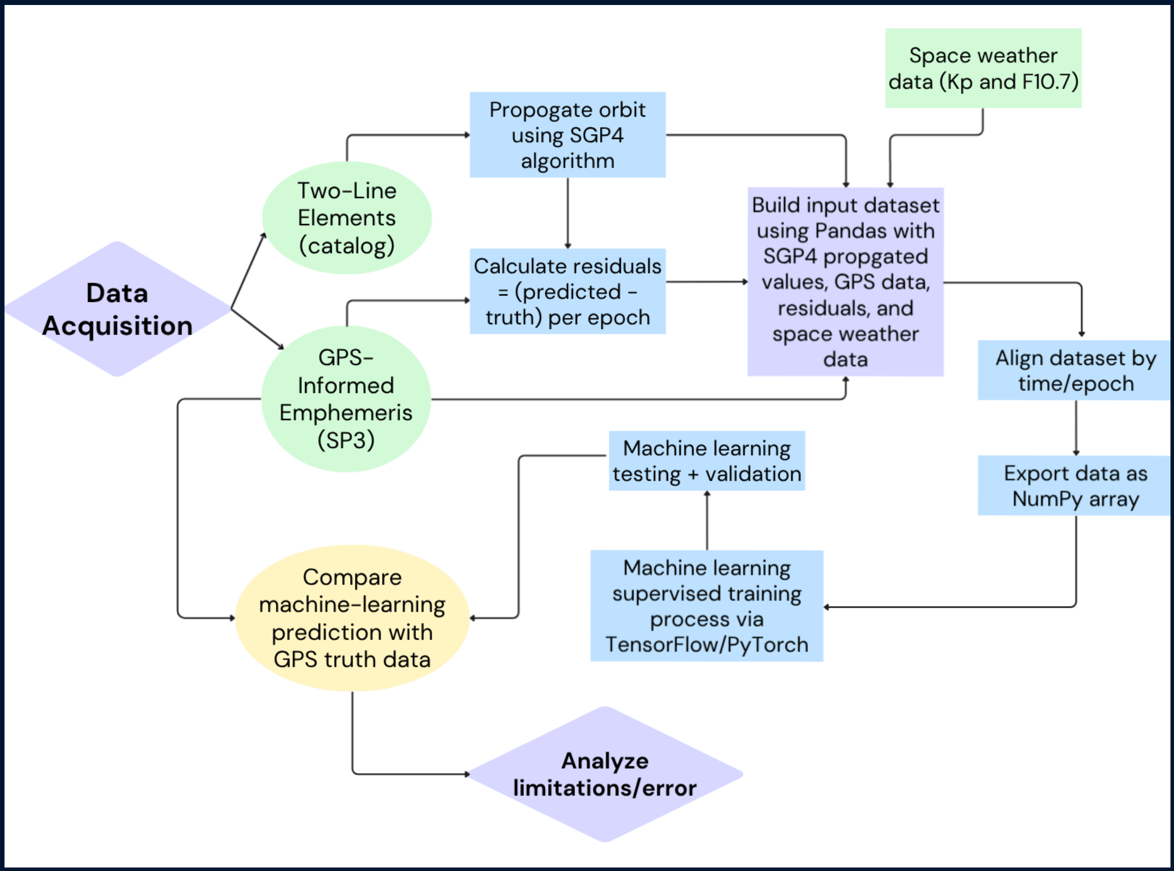


Fig. 3. Flowchart of proposed methodology for machine-learning enhanced orbital propagation, showing the basic steps of data acquisition, data preparation with residuals calculations, building an input data set, supervised machine-learning training, machine-learning validation, and analysis of error and limitations.

Figure 3 highlights how classical physics-based SGP4 predictions and GPS truth can be compared to find residual data, and when this data is enriched with space-weather indices, a supervised machine-learning model can be trained to output data about an object's predicted orbit. Comparing these outputs to an object's GPS truth data will provide information about the accuracy of the predictions, and, if they can improve the accuracy of SGP4 predictions, they may be used to enhance orbit propagation of thousands of objects in LEO space.

Significance

This project is significant because it tackles a major, relevant issue pertaining to LEO operations, specifically how orbit-prediction error grows with time when using physics-based propagation methods. An AI-augmented approach where a supervised machine learning model is trained on historical data pairs – TLE data that is SGP4 propagated compared with GPS truth data – to identify patterns and relationships can possibly help to extend the prediction times of orbits, and reduce inaccuracies in modern-day trajectory propagation. Without any extra equipment or data needing to be collected, the machine-learning algorithm will be fast, relatively inexpensive, and has the potential to generate outputs that can directly translate to more timely and accurate conjunction assessments to keep spacecraft safe in LEO space.

However, the broader impact of this project reaches far beyond just a single study. A reliable truth-anchored correction model can help worldwide organizations that produce and launch satellites to avoid unnecessary maneuvers in space due to unreliable orbit propagation, saving fuel and money. The consistency of space-safety decisions has the potential to be increased, which would allow for growing satellite constellations, moving the world a step towards better, more efficient communication, navigation, and scientific research stemming from satellites. The hybrid physics-based and machine-learning augmented system that will be produced in the project can also be transferred to other applications and fields, such as a correction model which could be used for space-weather prediction [9] and navigation in space. If successful in increasing the accuracy of orbit predictions of satellites and/or orbital debris, a new machine-learning model can be applied to countless other model-driven fields as well, and helping to improve space situational awareness for researchers and organizations.

References

[1] "Space debris by the numbers." Accessed: Oct. 14, 2025. [Online]. Available: https://www.esa.int/Space_Safety/Space_Debris/Space_debris_by_the_numbers

[2] H. G. Lewis, "Understanding long-term orbital debris population dynamics," *J. Space Saf. Eng.*, vol. 7, no. 3, pp. 164–170, Sep. 2020, doi: 10.1016/j.jsse.2020.06.006.

[3] P. By, "ESA'S ANNUAL SPACE ENVIRONMENT REPORT." [Online]. Available: https://www.sdo.esoc.esa.int/environment_report/Space_Environment_Report_latest.pdf

[4] H. Peng and X. Bai, "Improving orbit prediction accuracy through supervised machine learning," *arXiv [astro-ph.EP]*, Jan. 15, 2018. doi: 10.1016/j.asr.2018.03.001.

[5] D. Wright, "Colliding Satellites: Consequences and Implications," Feb. 26, 2009.

[6] C. Pardini and L. Anselmo, "Assessment of the consequences of the Fengyun-1C breakup in low Earth orbit," *Adv. Space Res.*, vol. 44, no. 5, pp. 545–557, Sep. 2009, doi: 10.1016/j.asr.2009.04.014.

[7] V. Suthakar, A. A. Sanvido, R. Qashoa, and R. S. K. Lee, "Comparative analysis of resident space object (RSO) detection methods," *Sensors (Basel)*, vol. 23, no. 24, Dec. 2023, doi: 10.3390/s23249668.

[8] B. Li, J. Huang, Y. Feng, F. Wang, and J. Sang, "A machine learning-based approach for improved orbit predictions of LEO space debris with sparse tracking data from a single station," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4253–4268, Dec. 2020, doi: 10.1109/TAES.2020.2989067.

[9] F. Caldas and C. Soares, "Precise and efficient Orbit Prediction in LEO with machine learning using exogenous variables," *arXiv [cs.LG]*, Jul. 03, 2024. [Online]. Available: http://arxiv.org/abs/2407.11026

[10] J. W. Allworth, "A Machine Learning Approach to Space Debris Characterisation and Classification using Ground Based Optical Observations," 2022. Accessed: Sep. 15, 2025. [Online]. Available: https://hdl.handle.net/2123/29185

[11] J. Alzubi, A. Nayyar, and A. Kumar, "Machine learning from theory to algorithms: An overview," *J. Phys. Conf. Ser.*, vol. 1142, no. 1, p. 012012, Nov. 2018, doi: 10.1088/1742-6596/1142/1/012012.

[12] "Closing the Gap Between SGP4 and High-Precision Propagation via Differentiable Programming." Accessed: Oct. 20, 2025. [Online]. Available: https://arxiv.org/html/2402.04830v3

[13] M. Kirschner *et al.*, "Orbit precision analysis of small man-made space objects in LEO based on radar tracking measurements," 2012.

[14] "Orbit GEO SP3 File – e-POP on CASSIOPE." Accessed: Oct. 19, 2025. [Online]. Available: https://epop.phys.ucalgary.ca/data-handbook/orbit-geo-sp3-file/

[15] A. Thakur, A. Kumar, and S. Awasthi, "Familiarization with Pandas for data science and analysis," *IRJMETS*, Apr. 2024, doi: 10.56726/IRJMETS55073.

[16] S. Prabu, P. Hancock, X. Xiang, and S. J. Tingay, "Demonstration of Orbit Determination for LEO Objects using the Murchison Widefield Array," *arXiv [astro-ph.IM]*, Aug. 08,

2023. [Online]. Available: http://arxiv.org/abs/2308.04640
[17] J. A. Gaebler, C. Wetterer, and K. Hill, “Cislunar Initial Orbit Determination with Optical Tracklets.”
[18] M. Jah, S. A. Huges, M. P. Wilkins, and T. Kelecy, “The general mission analysis tool (GMAT): A new resource for supporting debris orbit determination, tracking and analysis,” vol. 672, p. 124, Mar. 2009.
[19] V. Tyagi, “An Exploratory Data Analysis of Satellite Data,” Medium. Accessed: Oct. 20, 2025. [Online]. Available: https://medium.com/@varun.tyagi83/an-exploratory-data-analysis-of-satellite-data-fbc5b222144b
[20] M. Lara, J. F. San-Juan, L. M. López-Ochoa, and P. Cefola, “Long-term evolution of Galileo operational orbits by canonical perturbation theory,” *Acta Astronaut.*, vol. 94, no. 2, pp. 646–655, Feb. 2014, doi: 10.1016/j.actaastro.2013.09.008.
[21] D. Smilkov *et al.*, “TENSORFLOW.JS: MACHINE LEARNING FOR THE WEB AND BEYOND.”

🛰️ Introduction Section Paper

MACHINE-LEARNING ENHANCED ORBIT PROPAGATION: IMPROVING LOW EARTH ORBIT PREDICTION USING TLE AND GPS DATA

by

Elise Protti

Andrew Howarth, University of Calgary Department of Physics and Astronomy

Introduction Section Paper

Applied Science Project

December 15, 2025

The surge in the number of pieces of orbital debris in Earth's orbit has become an increasing problem ever since humanity started exploring and expanding into space at the start of the space age in 1957 [1], [2], launching commercial operators and large constellations of satellites to provide data about factors such as climate, atmospheric conditions, and navigation on Earth [3]. The accumulation of a large population of debris surrounding the Earth is making it extremely difficult to maintain sustainability in space in the long-run due to the dangers of space debris striking satellites and spacecraft. It is estimated that there are around 34,000 trackable objects that are 10 cm or larger in Earth's orbit, posing significant challenges to space travel and development in space [2].

Figure 1 illustrates the scale as well as composition of the orbital environment targeted by this innovative study, as reported by the European Space Agency's 2025 annual report [3]. Figure 1 is a useful graph showing statistics of growth of cataloged objects across all orbital classes over time. These objects are differentiated by object type– such as unidentified objects, rocket mission related objects, rocket debris, rocket fragmentation debris, rocky bodies, payload mission related objects, payload debris, payload fragmentation debris, and payloads – and the trend of the data illuminates that after decades of gradual increase in the number of objects in orbit, the curve steepens dramatically in the last decade as large constellations of satellites and major collisions add active spacecraft and debris, respectively [3].

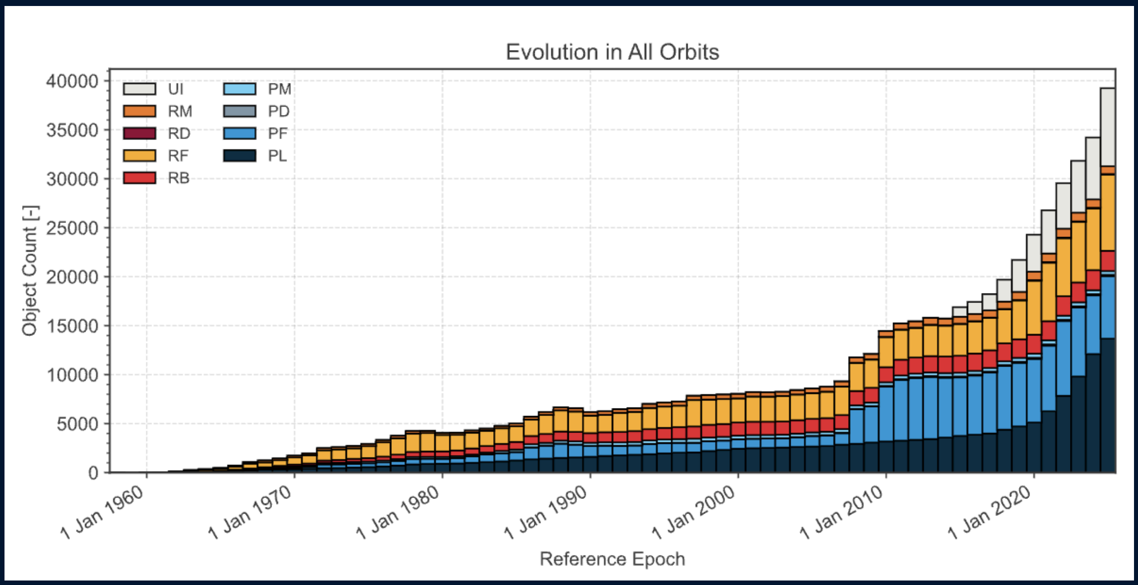


Fig. 1. Evolution of object counts in all of Earth's orbits stacked by object class from 1960 to present. Adapted from [3]. The objects are differentiated by object class – unidentified objects, rocket mission related objects, rocket debris, rocket fragmentation debris, rocky bodies, payload mission related objects, payload debris, payload fragmentation debris, and payloads, respectively – and show a gradual then steep increase in the total number of objects in Earth's orbit in the last decade.

The rapid increase in all object classes in Earth's orbit underscores the need for automated, machine learning-assisted predictions since orbit propagations are increasingly necessary for timely and accurate risk assessment in space. Due to orbital debris' inherent lack of navigational systems, in contrast to satellites and spacecraft, uncertainty arises about the exact trajectories of these objects, adding another level of risk to human-made spacecraft. Several incidents have occurred in history where satellites and spacecraft have collided with objects in space, in turn creating a larger, vast field of debris which continues the cycle of debris in Earth's orbit and increases the risk level even more [4]. An example of this was the February 10, 2009 collision of the U.S. Iridium 33 communication satellite with the derelict Russian Cosmos 2251 communication satellite both orbiting the Earth at an altitude of around 790 km [5], [6]. Their collision speed was about 11 km/s, and this event resulted in the creation of two new major debris fields that contributed to the dispersal of orbital debris around Low Earth Orbit (LEO) space [5], [6].

Figure 2 aids in visually showing how space debris clouds resulting from catastrophic collisions in LEO can spread and diffuse quickly, illustrating how the large amount of debris in Earth's orbit is a sizable risk to human spacecraft and satellites [5]. Therefore, the need for accurate and prompt orbit propagation systems to ensure all spacecraft stay intact is more crucial than ever in the modern age.

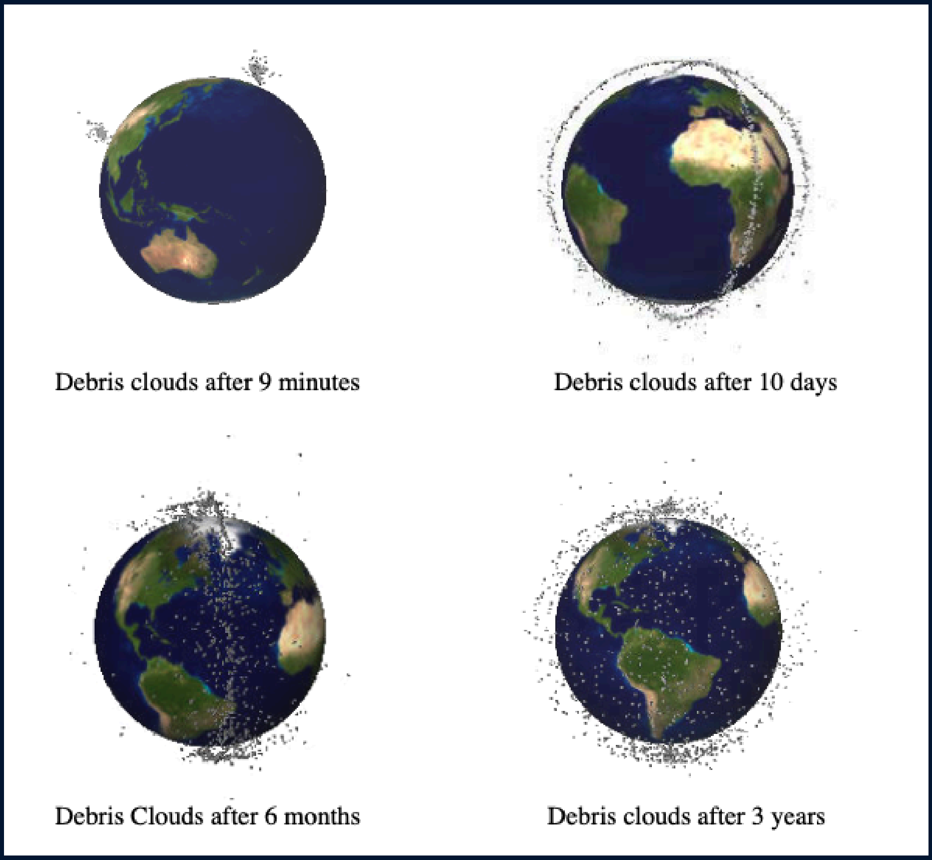


Fig. 2. Timestamped projection diagram showing the general pattern of the spread of debris clouds. Adapted from [5]. The debris clouds' distribution is imaged approximately 9 minutes, 10 days, 6 months, and 3 years after a major collision resulting in debris formation.

Obtaining accurate predictions of satellite trajectories through the use of detection algorithms is a crucial part of space-situational awareness (SSA) – which refers to the knowledge of specific locations of objects and hazards in space – and it ensures spacecraft and satellites are kept safe while in orbit [7]. Recent improvements in the sensor capabilities of space surveillance technologies have even made it possible to track and catalog smaller pieces of debris more than ever before [3], which opens up the opportunity for the orbits of these pieces of debris to be analyzed and predicted using orbit propagation techniques.

The current popular method of propagating and predicting orbital trajectories involves a physics-based method, which can produce volatile predictions due to many factors of space, such as atmospheric drag, solar radiation pressure, and Earth's gravity [4]. The surveillance equipment required to collect these precise measurements is very expensive and the process of data acquisition is time consuming, making it almost impossible to have necessary, timely orbital predictions [4]. Additionally, the errors from this type of physics-based orbital prediction, which typically uses Simplified General Perturbations-4 (SGP4) propagator systems, increase very rapidly as time goes on since over-simplifications are made by the system, especially when trying to predict the orbit of an object in LEO over extended periods of time or long distances [8], [9]. Therefore, physics-based orbital propagation is a pressing issue in the modern-day fields of aerospace development and satellite tracking since there is a very large potential for error, and a few meters in LEO space could make all the difference between a close-call encounter between pieces of debris or a major collision [8].

Machine learning (ML) methods offer a promising new approach to addressing this issue, as the predictions of orbits can be made without the explicit data points of atmospheric drag, gravitational forces, and more [4]. In contrast to physics-based orbital propagation methods, ML algorithms work similar to a human brain's neural networks by taking in large amounts of data, recognizing patterns, and using this past data to predict future events based on patterns [4].

Specifically, supervised learning is a ML training technique where labelled input data with corresponding known outputs is inserted into a ML model [4], [9]. This method of training is very effective since the algorithm is able to devise patterns and relationships within the data to propagate future orbits [4], [9].

A key data type in the ML process are Two-Line Element (TLE) sets. TLEs are an example of a public, open-source data record for resident space objects (RSOs), encoding an object's orbit at a specific epoch (moment in time), and can also be propagated with SGP4 to predict their an object's future position – but not without significant error growth over time [8], [10]. TLEs exist for thousands of cataloged objects, always appearing in a consistent format, making them an extremely useful, standardized type of data that can be sorted through easily by a ML algorithm [4]. Leveraging these consistent, widely available records of an object's position, paired with their SGP4 propagated states will offer excellent input data for a ML model since the model will be able to look for patterns and comparisons between an object's actual position and the position that SGP4 predicted.

Prior work in this field includes utilizing and developing supervised and reinforced ML algorithms as an alternative to physics-based propagation methods [4]. Most of these systems utilize known "truth" data – which is a piece of verified, true information about an object at a given time, such as its position or velocity – used in conjunction with ground-based data such as solar flux or geomagnetic index measurements [8], [10], to make predictions about the error level of physics-based SGP4-propagated data [4]. Data types historically used for this step have included simulation-based space catalog environments, RSO radar-based observational data, image-based data, and more [4], [10]. The accuracy and usefulness of these ML systems have depended on many factors which set each research study apart, such as: the type of ML algorithm used, the type of ML style used, the program or development environment used to implement it, the accuracy of the data inputted, and the time range examined [4], [8], [10]. Overall, though, many experiments have seen success in their ML models, with measurable reductions in orbit-prediction error after the ML-based corrections were implemented, although the magnitude of the improvement varies across different studies [4], [8], [10].

As one representative example, researchers Hao Peng and Xiaoli Bai [4] propose a hybrid orbital-prediction method in their 2018 study that uses a physics-based propagator, but also a supervised ML model to learn the prediction error from the historical "truth" data and physics-based propagation values. As part of their study, they applied that the ML "learned error" as a correction to the ML model's outputs in the hope for more accurate results with less error over time. They used a performance metric for the percentage of residual error after ML correction relative to the original error in the physics-based propagated states, and reported that this metric experienced strong reductions. They concluded overall that the supervised ML model they used could reduce prediction error by more than half across the many generalization tests they conducted, which supports the continued potential of developing ML-based correction approaches especially in the orbital object prediction field [4].

Different ML systems use different classifiers to approach data sets, such as random decision forests, which combine multiple decision "trees" to improve the accuracy and robustness of prediction [8], and support vector machines, which plot points in n-dimensional space based on the number of features in the dataset to come to conclusions [10], [11]. The differences in these methods can translate to the accuracy of the results given. In general, though, the prediction error of orbital trajectories has decreased due to the general usefulness and accuracy of ML-augmented predictions that help minimize the physics-based prediction errors

[4]. Since the ML field and information about the increasing amounts of orbital debris in LEO space is relatively new, many scientists hint that more research needs to be done and more ML algorithms need to be tested for indisputable conclusions to be made on the subject area [4]. Specifically, in the analysis of the aforementioned 2018 study, Peng and Bai wrote that "further studies are required to draw concrete conclusions" about the effectiveness of supervised ML models and the accuracy of their orbital position outputs [4].

Therefore, this innovative study is aimed at specifically targeting the error growth of physics-based orbit propagation methods. The study will focus on training a supervised ML model (with a neural-network baseline) using satellites that provide precise "truth" orbits via GPS data in comparison to their predicted future positions propagated by a standard SGP4 algorithm. By systematically learning, the ML model will be able to identify the residual error between the satellites' predicted and observed states, and can possibly increase the accuracy of orbit propagation by relying on patterns in the data to augment physics-based orbit propagation methods, which prove to be inaccurate over long periods of time or distance [12]. This same algorithm may also be applied to debris objects that are forced to rely only on cataloged data, as they do not have the ability to have their exact position "truths" determined. However, it is important to note that data gathered from radar-based techniques on ground observatories will never be as accurate as GPS data on the satellite itself. GPS data is generally accurate to decimeter or centimeter levels, whereas radar-based data is only accurate to a scale of meters because of the nature of the data collection process [13]. Because orbit predictions are essential in orbital operational decision-making, this study, if successful, has the potential to reap significant positive impacts on the space and astrophysics industries, and will help to maintain the safety of satellites in space by taking a step in decreasing collision frequency.

A secondary focus for this study is examining whether ML model performance depends on the type of space object being analyzed and its orbital characteristics. This study will use data from the University of Calgary's CASSIOPE satellite [14] and one of the European Space Agency's four Swarm satellites [15].

The CASSIOPE satellite, fully developed and operated by the University of Calgary, is a polar-orbiting satellite in an elliptical LEO orbit that provides high-accuracy GPS-based orbit data suitable to use as "truth" data in residual calculations due to its array of five dual-frequency GPS receivers [14]. These GPS receivers are used to derive the satellite's position and velocity with extreme accuracy, with up to 100 GPS samples per second, meaning CASSIOPE's GPS data is very accurate and reliable, and an excellent match for the study's goal of using an ML model to correct physics-based SGP4 propagation error growth over time [14]. Additionally, ESA's near-polar Swarm satellites also use advanced onboard GPS dual-frequency receivers for precise orbit determination with near-continuous coverage, enabling accuracy down to the centimeter level [15]. Swarm's primary mission objective is to study the dynamics of Earth's magnetic field and how the magnetic field interacts with the Earth. All four Swarm satellites carry advanced instruments like magnetometers, a tool used to measure the strength and direction, as well as changes to Earth's magnetic fields [15]. Also, the Swarm satellites are equipped with an accelerometer that is used to derive thermospheric density and winds as a secondary mission objective [15]. Because both of these satellites support scientific objectives of high-value and further scientific development, improving multi-day orbit predictions can reduce the collision risk and increase the safety of these satellites in the increasingly dangerous LEO environment.

In all, the overall, general objective is to determine, by using a systematic and quantitative comparison, whether a supervised ML correction model can meaningfully reduce extended day LEO propagation error relative to the SGP4 propagations that are currently in widespread use in the space industry now [8]. If successful, the results from this study would make space situation awareness more reliable by improving the accuracy of orbital predictions as well as reducing the unnecessary workload from avoidable false alerts due to the inaccuracies of traditional physics-based propagations [9].

References

[1] "Space debris by the numbers." Accessed: Oct. 14, 2025. [Online]. Available: https://www.esa.int/Space_Safety/Space_Debris/Space_debris_by_the_numbers

[2] H. G. Lewis, "Understanding long-term orbital debris population dynamics," *J. Space Saf. Eng.*, vol. 7, no. 3, pp. 164–170, Sep. 2020, doi: 10.1016/j.jsse.2020.06.006.

[3] P. By, "ESA'S ANNUAL SPACE ENVIRONMENT REPORT." [Online]. Available: https://www.sdo.esoc.esa.int/environment_report/Space_Environment_Report_latest.pdf

[4] H. Peng and X. Bai, "Improving orbit prediction accuracy through supervised machine learning," *arXiv [astro-ph.EP]*, Jan. 15, 2018. doi: 10.1016/j.asr.2018.03.001.

[5] D. Wright, "Colliding Satellites: Consequences and Implications," Feb. 26, 2009.

[6] C. Pardini and L. Anselmo, "Assessment of the consequences of the Fengyun-1C breakup in low Earth orbit," *Adv. Space Res.*, vol. 44, no. 5, pp. 545–557, Sep. 2009, doi: 10.1016/j.asr.2009.04.014.

[7] V. Suthakar, A. A. Sanvido, R. Qashoa, and R. S. K. Lee, "Comparative analysis of resident space object (RSO) detection methods," *Sensors (Basel)*, vol. 23, no. 24, Dec. 2023, doi: 10.3390/s23249668.

[8] B. Li, J. Huang, Y. Feng, F. Wang, and J. Sang, "A machine learning-based approach for improved orbit predictions of LEO space debris with sparse tracking data from a single station," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4253–4268, Dec. 2020, doi: 10.1109/TAES.2020.2989067.

[9] F. Caldas and C. Soares, "Precise and efficient Orbit Prediction in LEO with machine learning using exogenous variables," *arXiv [cs.LG]*, Jul. 03, 2024. [Online]. Available: http://arxiv.org/abs/2407.11026

[10] J. W. Allworth, "A Machine Learning Approach to Space Debris Characterisation and Classification using Ground Based Optical Observations," 2022. Accessed: Sep. 15, 2025. [Online]. Available: https://hdl.handle.net/2123/29185

[11] J. Alzubi, A. Nayyar, and A. Kumar, "Machine learning from theory to algorithms: An overview," *J. Phys. Conf. Ser.*, vol. 1142, no. 1, p. 012012, Nov. 2018, doi: 10.1088/1742-6596/1142/1/012012.

[12] "Closing the Gap Between SGP4 and High-Precision Propagation via Differentiable Programming." Accessed: Oct. 20, 2025. [Online]. Available: https://arxiv.org/html/2402.04830v3

[13] M. Kirschner *et al.*, "Orbit precision analysis of small man-made space objects in LEO based on radar tracking measurements," 2012.

[14] A. W. Yau and H. G. James, "CASSIOPE enhanced polar outflow probe (e-POP) mission overview," *Space Sci. Rev.*, vol. 189, no. 1–4, pp. 3–14, Jun. 2015, doi: 10.1007/s11214-015-0135-1.

[15] J. van den IJssel, B. Forte, and O. Montenbruck, "Impact of Swarm GPS receiver updates on POD performance," *Earth Planets Space*, vol. 68, no. 1, Dec. 2016, doi: 10.1186/s40623-016-0459-4.

🛰️ Methodology Section Paper

MACHINE-LEARNING ENHANCED ORBIT PROPAGATION: IMPROVING LOW EARTH ORBIT PREDICTION USING TLE AND GPS DATA

by

Elise Protti

Andrew Howarth, University of Calgary Department of Physics and Astronomy

Methodology Section Paper

Applied Science Project

December 15, 2025

*****Note: Methodology Section Paper is still in the drafting phase and is NOT YET COMPLETED**

I. Introduction

The surge in the number of pieces of orbital debris in Earth's orbit has become an increasing problem ever since humanity started exploring and expanding into space at the start of the space age in 1957 [1], [2], launching commercial operators and large constellations of satellites to provide data about factors such as climate, atmospheric conditions, and navigation on Earth [3]. The accumulation of a large population of debris surrounding the Earth is making it extremely difficult to maintain sustainability in space in the long-run due to the dangers of space debris striking satellites and spacecraft. It is estimated that there are around 34,000 trackable objects that are 10 cm or larger in Earth's orbit, posing significant challenges to space travel and development in space [2].

Figure 1 illustrates the scale as well as composition of the orbital environment targeted by this innovative study, as reported by the European Space Agency's 2025 annual report [3]. Figure 1 is a useful graph showing statistics of growth of cataloged objects across all orbital classes over time. These objects are differentiated by object type– such as unidentified objects, rocket mission related objects, rocket debris, rocket fragmentation debris, rocky bodies, payload mission related objects, payload debris, payload fragmentation debris, and payloads – and the trend of the data illuminates that after decades of gradual increase in the number of objects in orbit, the curve steepens dramatically in the last decade as large constellations of satellites and major collisions add active spacecraft and debris, respectively [3].

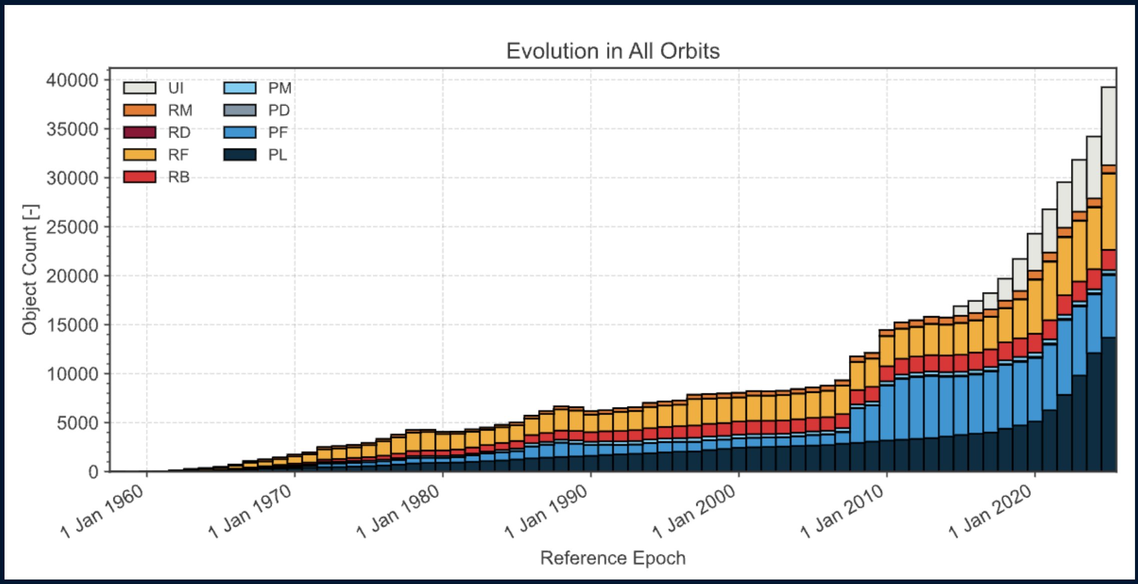


Fig. 1. Evolution of object counts in all of Earth's orbits stacked by object class from 1960 to present. Adapted from [3]. The objects are differentiated by object class – unidentified objects (gray), rocket mission related objects (orange), rocket debris (maroon), rocket fragmentation debris (yellow), rocky bodies (red), payload mission related objects (cyan), payload debris (dark gray), payload fragmentation debris (blue), and payloads (navy), respectively – and show a gradual then steep increase in the total number of objects in Earth's orbit in the last decade.

The rapid increase in all object classes in Earth's orbit underscores the need for automated, machine learning-assisted predictions since orbit propagations are increasingly necessary for timely and accurate risk assessment in space. Due to orbital debris' inherent lack of navigational systems, in contrast to satellites and spacecraft, uncertainty arises about the exact trajectories of these objects, adding another level of risk to human-made spacecraft. Several incidents have occurred in history where satellites and spacecraft have collided with objects in space, in turn creating a larger, vast field of debris which continues the cycle of debris in Earth's orbit and increases the risk level even more [4]. An example of this was the February 10, 2009 collision of the U.S. Iridium 33 communication satellite with the derelict Russian Cosmos 2251 communication satellite both orbiting the Earth at an altitude of around 790 km [5], [6]. Their collision speed was about 11 km/s, and this event resulted in the creation of two new major debris fields that contributed to the dispersal of orbital debris around Low Earth Orbit (LEO) space [5], [6].

Figure 2 illustrates how a large amount of space debris clouds resulting from catastrophic collisions in LEO can spread and diffuse quickly, illustrating how the large amount of debris in Earth's orbit is a sizable risk to human spacecraft and satellites [5]. Therefore, the need for accurate and prompt orbit propagation systems to ensure all spacecraft stay intact is more crucial than ever in the modern age.

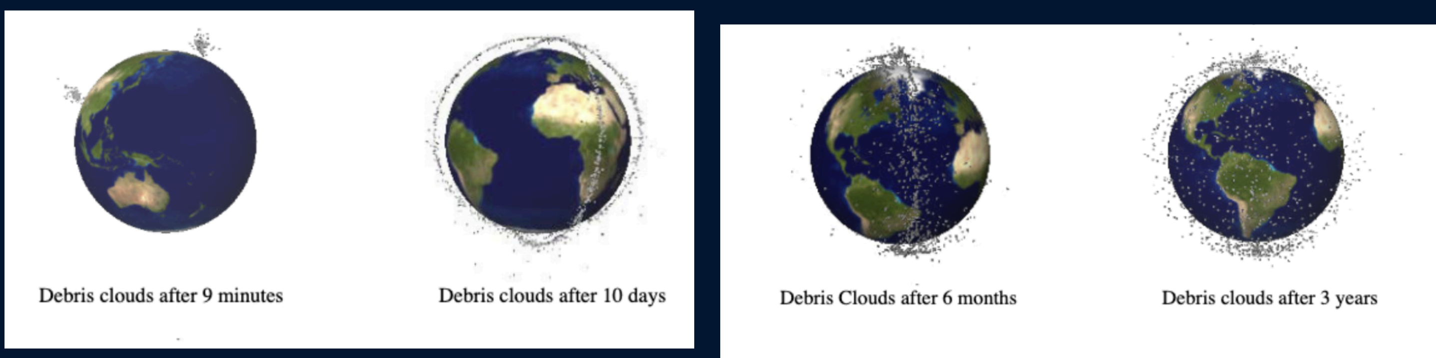


Fig. 2. Timestamped projection diagram showing the general pattern of the spread of debris clouds. Adapted from [5]. The debris clouds' distribution is imaged approximately 9 minutes, 10 days, 6 months, and 3 years (left to right) after a major collision resulting in debris formation.

Obtaining accurate predictions of satellite trajectories through the use of detection algorithms is an essential part of space-situational awareness (SSA), and it ensures spacecraft and satellites are kept safe while in orbit [7]. Recent improvements in the sensor capabilities of space surveillance technologies have even made it possible to track and catalog smaller pieces of debris more than ever before [3], which opens up the opportunity for the orbits of these pieces of debris to be analyzed and predicted using orbit propagation techniques.

The current popular method of propagating and predicting orbital trajectories involves a physics-based method, which can produce volatile predictions due to many factors of space, such as atmospheric drag, solar radiation pressure, and Earth’s gravity [4]. The surveillance equipment required to collect these precise measurements is expensive, and data acquisition is time consuming, making it almost impossible to have necessary, timely orbital predictions [4]. Additionally, the errors from this type of physics-based orbital prediction, which typically uses Simplified General Perturbations-4 (SGP4) propagator systems, increase very rapidly as time goes on since over-simplifications are made by the system, especially when trying to predict the orbit of an object in LEO over extended periods of time or long distances [8], [9]. Therefore, physics-based orbital propagation is a pressing issue in the modern-day fields of aerospace development and satellite tracking since there is a very large potential for error, and a few meters in LEO space could make all the difference between a close-call encounter between pieces of debris or a major collision [8].

Machine learning (ML) methods offer a promising new approach to addressing this issue, as the predictions of orbits can be made without the explicit data points of atmospheric drag, gravitational forces, among other factors [4]. In contrast to physics-based orbital propagation methods, ML algorithms work similar to a human brain's neural networks by taking in large amounts of data, recognizing patterns, and using this past data to predict future events based on patterns [4]. Specifically, supervised learning is a ML training technique where labelled input data with corresponding known outputs is inserted into a ML model [4], [9]. This method of training is very effective since the algorithm is able to devise patterns and relationships within the data to propagate future orbits [4], [9].

A key data type in the ML process are Two-Line Element (TLE) sets. TLEs are an example of a public, open-source data record for resident space objects (RSOs), encoding an object's orbit at a specific epoch. However, when TLEs are propagated with SGP4 to predict an object's future position, the resulting positions will have significant error growth when compared to the true position of the object [8], [10]. TLEs exist for thousands of cataloged objects, always appearing in a consistent format, making them an extremely useful, standardized type of data that can be sorted through easily by a ML algorithm [4]. Leveraging these consistent, widely available records of an object's position, paired with their SGP4 propagated states will offer excellent input data for a ML model since the model will be able to look for patterns and comparisons between an object’s actual position and the position that SGP4 predicted.

Prior work in this field includes utilizing and developing supervised and reinforced ML algorithms as an alternative to physics-based propagation methods [4]. Most of these systems utilize known “truth” data such as its position or velocity in conjunction with ground-based data like solar flux or geomagnetic index measurements [8], [10] to make predictions about the error level of physics-based SGP4-propagated data [4]. Data types historically used for this step have included simulation-based space catalog environments, RSO radar-based observational data, image-based data, among others [4], [10]. The accuracy and usefulness of these ML systems have depended on many factors such as: the type of ML algorithm used, the type of ML style used, the program or development environment used to implement it, the accuracy of the data inputted, and the time range examined [4], [8], [10]. Overall, many experiments have been successful in their ML models, with measurable reductions in orbit-prediction error after the ML-based corrections were implemented, although the magnitude of the improvement varies across different studies [4], [8], [10].

Different ML systems use different classifiers to approach data sets, such as random decision forests, which combine multiple decision “trees” to improve the accuracy and robustness of prediction [8], and support vector machines, which plot points in n-dimensional space based on the number of features in the dataset to come to conclusions [10], [11]. The differences in these methods can translate to the accuracy of the results given. In general, the prediction error of orbital trajectories has decreased due to the general usefulness and accuracy of ML-augmented predictions that help minimize the physics-based prediction errors [4]. Since the ML field and information about the increasing amounts of orbital debris in LEO space is relatively new, many scientists hint that more research needs to be done and more ML algorithms need to be tested for indisputable conclusions to be made on the subject area [4].

As one representative example, Peng and Bai [4] propose a hybrid orbital-prediction method in their 2018 study that uses a physics-based propagator, but also a supervised ML model to learn the prediction error from the historical "truth" data and physics-based propagation values. As part of their study, they applied the ML "learned error" as a correction to the model's outputs in the hope for more accurate results with less error over time. A performance metric was reported for the percentage of residual error after ML correction relative to the original error in the physics-based propagated states. This metric was reported to have experienced strong reductions, meaning the residual error values after the ML correction were lower than they were before the model's correction. Overall, the supervised ML model they used could reduce prediction error by more than half across the many generalization tests they conducted, which supports the continued potential of developing ML-based correction approaches especially in the orbital object prediction field [4]. In their analysis, they also wrote that "further studies are required to draw concrete conclusions" about the effectiveness of supervised ML models and the accuracy of their orbital position outputs [4].

FIX LOCATION. The CASSIOPE satellite, fully developed and operated by the University of Calgary, is a polar-orbiting satellite in an elliptical LEO orbit that provides high-accuracy GPS-based orbit data suitable to use as "truth" data in residual calculations due to its array of five dual-frequency GPS receivers [12]. These GPS receivers are used to derive the satellite's position and velocity with extreme accuracy, with up to 100 GPS samples per second, meaning CASSIOPE's GPS data is very accurate and reliable, and an excellent match for the study's goal of using an ML model to correct physics-based SGP4 propagation error growth over time [12]. Additionally, ESA's near-polar Swarm satellites also use advanced onboard GPS dual-frequency receivers for precise orbit determination with near-continuous coverage, enabling accuracy down to the centimeter level [13]. Swarm's primary mission objective is to study the dynamics of Earth's magnetic field and how the magnetic field interacts with the Earth. All four Swarm satellites carry advanced instruments like magnetometers, a tool used to measure the strength and direction, as well as changes to Earth's magnetic fields [13]. Also, the Swarm satellites are equipped with an accelerometer that is used to derive thermospheric density and winds as a secondary mission objective [13]. Because both of these satellites support scientific objectives of high-value and further scientific development, improving multi-day orbit predictions can reduce the collision risk and increase the safety of these satellites in the increasingly dangerous LEO environment.

This study aims at specifically targeting the error growth of physics-based orbit propagation methods. A supervised ML model (with a neural-network baseline) using satellites that provide precise "truth" orbits via GPS data in comparison to their predicted future positions propagated by a standard SGP4 algorithm. The ML model characterizes the residual errors between predicted and observed satellite states, thereby enhancing the accuracy of orbit propagation. This approach augments existing physics-based methods, which prove to be inaccurate over long periods of time or distance [14]. FIX LOCATION. This same algorithm may also be applied to debris objects that are forced to rely only on cataloged data, as they do not have the ability to have their exact position “truths” determined. However, it is important to note that data gathered from radar-based techniques on ground observatories will never be as accurate as GPS data on the satellite itself. GPS data is generally accurate to decimeter or centimeter levels, whereas radar-based data is only accurate to a scale of meters because of the nature of the data collection process [15]. Because orbit predictions are essential in orbital operational decision-making, this study, if successful, has the potential to reap significant positive impacts on the space and astrophysics industries, and will help to maintain the safety of satellites in space by taking a step in decreasing collision frequency.

A secondary objective is to investigate the sensitivity to the ML model to the specific physical and orbital characteristics of the objects under analysis. Data from the University of Calgary’s CASSIOPE satellite [12]and one of the European Space Agency’s four Swarm satellites is analyzed [13].

In all, the overall, general objective is to determine, by using a systematic and quantitative comparison, whether a supervised ML correction model can meaningfully reduce extended day LEO propagation error relative to the SGP4 propagations that are currently in widespread use in the space industry now [8]. If successful, the results from this study would make space situation awareness more reliable by improving the accuracy of orbital predictions as well as reducing the unnecessary workload from avoidable false alerts due to the inaccuracies of traditional physics-based propagations [9].

II. Methodology

A. Data Collection and Organization

This project uses archived 2019 data for the University of Calgary's CASSIOPE satellite currently in LEO space [2] , and combines GPS-derived "truth" ephemerides, TLE data for baseline propagation for the spacecraft and space weather indices used as other ML inputs.

1) *Truth Ephemerides:* Truth orbit states for the CASSIOPE satellite were obtained using the Orbit Geo SP3 product associated with the CASSIOPE mission data system. These truth states provide specific time-stamped position and velocity data in the International Terrestrial Reference Frame (ITRF) and in GPS time [3]. Truth ephemeris files from CASSIOPE's Orbit Geo SP3 product were downloaded for a six-month period spanning from January 1, 2019 to June 30, 2019.
2) *TLE Data:* For the same time period, CASSIOPE TLEs were collected as well as stored as the baseline inputs for SGP4 orbit propagation.
3) *Space Weather Data:* Space weather indices were included in the ML training since geomagnetic and solar activity have the ability to influence upper-atmosphere and LEO conditions, and affect drag-related prediction errors. The two space weather indices used in this project were 10.7 cm solar radio flux (F10.7), which is an indicator of solar activity. Additionally, the planetary K-index (Kp) was used as an indicator of geomagnetic disturbance level. These indices were joined and time-aligned to the orbit dataset features.

All data used for ML training and evaluation were previously collected and stored in open-source data sets, and the analysis was performed offline using archived data points rather than real-time onboard processing techniques. All data files used were organized by date and processed using Python scripts carried out in Visual Studio Code. The intermediate tables were stored as Pandas DataFrames to keep the steps reproducible for potential future use for other satellites' information.

B. Data Preparation, Parsing, and Cleaning

After the collection of all of the data for the CASSIOPE satellite, each data source was converted from raw text files and combined into one large synchronized dataset sorted by time.

1) *Parsing:* The SP3 truth ephemeris files were parsed using Python to programmatically read the files, extracted the time stamps and position/velocity state vectors for CASSIOPE at each recorded epoch, and converted them into a structured Pandas DataFrame (CITE). Data was stored in this tabular format to ensure that data could be joined by consistent timestamps.
2) *Data Cleaning:* Data points were excluded from the ephemeris if any of the following conditions occurred: (i) missing SP3 truth values at required timestamps, (ii) missing SGP4-propagated state (position and velocity) vectors at a specific timestamp, (iii) invalid timestamps or invalid numerical values in the DataFrame. These exclusion rules were applied consistently throughout the full dataset to ensure that the data was standardized and could be used for computations and ML training easily.
3) *Time Standardization:* SP3 time stamps, which represent the actual GPS time, were converted into a consistent data and time format inside of the Pandas DataFrame so that the SGP4 prediction times could be aligned seamlessly. Additionally, the F10.7 and Kp times were also aligned to this same time standardization format. However, since Kp is reported over intervals or multiple hours where F10.7 data values are typically reported

daily, the space weather values were aligned by using the most recent available value so that each epoch had a defined space weather value without any gaps.

C. Baseline Orbit Propagation and Residual Calculations

Once all the data in the dataset was cleaned and time-aligned, the baseline propagations were generated using SGP4 and converted over to a comparable reference frame so that the difference between the propagated states and the GPS truth data states could be computed accurately.

1) *SGP4 Propagation:* TLEs for CASSIOPE were grouped into weekly “packets” across the study period, producing 26 TLE packet sets. To create each packet, a selected TLE was propagated forward using SGP4 over a fixed two-week propagation window, and the resulting state vectors were generated at the same epochs as the SP3 truth samples to allow for easy comparison. This SGP4 output served as the baseline prediction dataset used to calculate residual error, and to build the training inputs for the supervised ML model.
2) *Reference-Frame Harmonization:* Since the SP3 truth states for CASSIOPE were provided in ITRF, the baseline SGP4 outputs were transformed into the same terrestrial frame before the residual error computation? (DO I NEED TO EXPLAIN HOW I DID THIS -> MENTION IT WAS DONE IN PYTHON?) This need for consistent reference frame systems is also stated in the e-POP documentation for the CASSIOPE satellite (CITE).
3) *Residual Error Calculation:* Residuals were computed as the difference between the baseline predicted position, using SGP4 propagation, and the truth position, given by the

CASSIOPE SP3 data files, at each epoch. A scalar 3D position error magnitude was also computed during this process (DO I NEED TO MENTION PYTHON SPECIFICS OF HOW THIS WAS DONE?) In order for residual error calculations to be accurate, both the baseline SGP4 propagated states and the CASSIOPE truth data were compared at identical epochs and also were compared in the same coordinate reference frame.

D. *Machine-Learning Correction*

The ML component of the project was implemented in Python using TensorFlow, which provides a framework for defining, training, and validating neural networks on large, structured datasets [16]. TensorFlow's system and design and training capabilities are described in [16]. The goal of the model was not to replace the SGP4 propagations, but to use SGP4 as the baseline propagator and then adjust the propagation inputs. The ML model was utilized like this so that the final propagated states could better match the SP3 truth trajectory of CASSIOPE over longer prediction times, which is when large residual errors occur. For every timestamp/epoch used in the dataset, a baseline state was producing using SGP4 propagation and marched to the corresponding truth state from the SP3 product at matching times. The difference between the vaseline prediction and the truth state at the samep epoch was treated as the residual error that the ML model needed to learn. This means the training data contained a consistent pair of information at each epoch: what SGP4 prediscts and what the SP3 truth product reports. The ML learning objective was based on this direct relationship between baseline and truth data for a given epoch. The primary correction strategy utilized in the study was a parameter-adjustment approach.

1) *Machine-Learning Model Inputs:* The neural network machine-learning model's input data set contained the time from the TLE epoch, the baseline position and velocity information at each timestamp, as well as Kp and F10.7 space weather indices.
2) Parameter-Adjustment Output: Rather than predicting a final position of the CASSIOPE satellite directly, the neural network output was instead a set of adjustment values to be applied to the propagation inputs. After the ML model produced these adjustments, a second SGP4 propagation was performed using the adjusted inputs to generate the corrected trajectory. This approach preserves the original operational structure of SGP4, which allowing for the learned adjustments to reduce the baseline residual error patterns.

E. Machine-Learning Evaluation and Reporting

To ensure that changes in orbit-prediction accuracy were attributable to the ML collection model rather than inconsistent processing, the following factors were held constant across all runs: (i) the target satellite (CASSIOPE), (ii) the truth data type and International Terrestrial Reference Frame (ITRF), (iii) the propagation system used (SGP4) and its implementation, (iv) the propagation window length per TLE packet (2 weeks), (v)

Baseline SGP4 Orbit Propagation

Residual Error Computation

Machine Learning Residual Error Correction

1) *Machine Learning Training*

2) *Machine Learning Validation*

Evaluation of Results and Statistical Analysis

References

[1] "Space debris by the numbers." Accessed: Oct. 14, 2025. [Online]. Available: https://www.esa.int/Space_Safety/Space_Debris/Space_debris_by_the_numbers

[2] H. G. Lewis, "Understanding long-term orbital debris population dynamics," *J. Space Saf. Eng.*, vol. 7, no. 3, pp. 164–170, Sep. 2020, doi: 10.1016/j.jsse.2020.06.006.

[3] P. By, "ESA'S ANNUAL SPACE ENVIRONMENT REPORT." [Online]. Available: https://www.sdo.esoc.esa.int/environment_report/Space_Environment_Report_latest.pdf

[4] H. Peng and X. Bai, "Improving orbit prediction accuracy through supervised machine learning," *arXiv [astro-ph.EP]*, Jan. 15, 2018. doi: 10.1016/j.asr.2018.03.001.

[5] D. Wright, "Colliding Satellites: Consequences and Implications," Feb. 26, 2009.

[6] C. Pardini and L. Anselmo, "Assessment of the consequences of the Fengyun-1C breakup in low Earth orbit," *Adv. Space Res.*, vol. 44, no. 5, pp. 545–557, Sep. 2009, doi: 10.1016/j.asr.2009.04.014.

[7] V. Suthakar, A. A. Sanvido, R. Qashoa, and R. S. K. Lee, "Comparative analysis of resident space object (RSO) detection methods," *Sensors (Basel)*, vol. 23, no. 24, Dec. 2023, doi: 10.3390/s23249668.

[8] B. Li, J. Huang, Y. Feng, F. Wang, and J. Sang, "A machine learning-based approach for improved orbit predictions of LEO space debris with sparse tracking data from a single station," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4253–4268, Dec. 2020, doi: 10.1109/TAES.2020.2989067.

[9] F. Caldas and C. Soares, "Precise and efficient Orbit Prediction in LEO with machine learning using exogenous variables," *arXiv [cs.LG]*, Jul. 03, 2024. [Online]. Available: http://arxiv.org/abs/2407.11026

[10] J. W. Allworth, "A Machine Learning Approach to Space Debris Characterisation and

Classification using Ground Based Optical Observations," 2022. Accessed: Sep. 15, 2025. [Online]. Available: https://hdl.handle.net/2123/29185

[11] J. Alzubi, A. Nayyar, and A. Kumar, "Machine learning from theory to algorithms: An overview," *J. Phys. Conf. Ser.*, vol. 1142, no. 1, p. 012012, Nov. 2018, doi: 10.1088/1742-6596/1142/1/012012.

[12] A. W. Yau and H. G. James, "CASSIOPE enhanced polar outflow probe (e-POP) mission overview," *Space Sci. Rev.*, vol. 189, no. 1–4, pp. 3–14, Jun. 2015, doi: 10.1007/s11214-015-0135-1.

[13] J. van den IJssel, B. Forte, and O. Montenbruck, "Impact of Swarm GPS receiver updates on POD performance," *Earth Planets Space*, vol. 68, no. 1, Dec. 2016, doi: 10.1186/s40623-016-0459-4.

[14] "Closing the Gap Between SGP4 and High-Precision Propagation via Differentiable Programming." Accessed: Oct. 20, 2025. [Online]. Available: https://arxiv.org/html/2402.04830v3

[15] M. Kirschner *et al.*, "Orbit precision analysis of small man-made space objects in LEO based on radar tracking measurements," 2012.

[16] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning".

🎤 Oral Presentation #1 - Research Proposal

LINK TO PRESENTATION:
ASP Research Proposal Oral Presentation - Elise Protti

PLANNING:
- Suggested order of slides
    - **Title (1 slide) w/ "punchline"**
    - **Intro (3/4 slides)**
    - **RQ/Goals/Objectives (2/3 slides)**
    - **Methodology (3/4 slide)**
    - **Significance/Conclusion (1 slide)**
    - **Conclusion/Ending slide (1 slide)**

Presentation #1: Lara

Notes:
- Assessing the adolescents have
- Contraceptives example:
    - Condoms, cervical cap/diaphragm, birth control pill
    - 80% experience negative menstruation effects
    - Hormonal contraceptives reduce blood loss, making menstruation less painful, and positive impacts
    - 2 types of IUDs (copper -> non-hormonal) and LNG-IUD (a hormonal IUD)
    - Adolescents have bad menstrual conditions and use less effective contraception -> RQ: what do adolescents know about IUDs, and where did they get this info?
- Methodology:
    - Survey -> will have anonymity in the survey so you do not have to give name, DOB, etc.
    - Data Collection and Analysis
    - Content Audit
    - Data Analysis
    - Distribute Findings

Question(s):
- If you have thought about the long-run, where will your results be distributed and what is the end goal for your project? Where will this data be used?

    - She wants to give the data to hospitals and specifically doctors so that the information can be widespread and patients will know that LNG-IUDs are an option for them

Presentation #2: Kinjal

Notes:

- Can POTS symptoms be measured while using LBNP?
    - POTS = increase in heart rate by 30bpm or more upon standing -> restricted blood flow, blood pools to lower regions of the heart when POTS patients stand up
    - Symptoms of POTS:
        - Nausea, rapid heart beat, brain fog, chest pain
        - Hemodynamic and Parasympathetic activity
    - Processes:
        - Valsalva Maneuver (measure of baroreflex function)
        - Sinus Arrhythmia: variation in heart rate assesses parasympathetic activity
        - Hyperventilation: peak wheat rate and blood pressure are compared
- LBNP = Lower Body Negative Pressure -> simulates gravity, so blood moves like how it would when a patient stands up
- Independent variable: With LBNP and without LBNP
- Responding variable: Hemodynamic responses and parasympathetic activity
- Controlled variables: all patients are POTS patients, measuring impacts with LBNP
- POTS patients only feel the effects of POTS when they are standing -> can simulate POTS conditions with standing laying down/supine

Question(s):

- What percentage of the population is affected by POTS -> significance?
- How are POTS patients affected in their everyday lives?
    - It leads to patient feeling nauseous every time they stand up and makes everyday life very hard

Presentation #3: Shaayaan

Notes:

- Develop a material/coating that can protect spacecraft in space from high temps and erosion
- HEA (high-entropy alloys) contain 3-6 metals in equal proportions -> can get characteristics from many elements instead of just one
- Current one-element alloys have the characteristics of just on element -> can be a disadvantage
- Superalloys perform well at high temps but fail under stress
    - HEAs can offer improved durability and performance
- His HEA will be using Cr-Fe-Co ratios
- Independent variable: Cr-Fe-Co composition, cold spray parameters, gas pressure
- Dependent variables: oxidation resistance oxidation resistance, adhesion strength, and strength integrity
- Powder Metallurgy flattens and can weld different elements together
- Types of testing:
    - Scratch tests to simulate debris impact
    - Thermal cycling
    - Expose coatings to high temp -> look for weight change
    - SEM to detect cracks
- An effective HEA reduces maintenance costs and enhances safety and reliability of aerospace components

Question(s):

- Why are you using Cr Fe and Co for your high entropy alloy? Which of their properties are desirable?
- How do you measure oxidation resistance, adhesion strength, and strength integrity?
- How are you going to test for high temp conditions?
    - Hard to simulate, put can put the substrate and the coating in a chamber that is heated up -> will be testing until about 400 degrees celsius

Presentation #4: Samir

Notes:

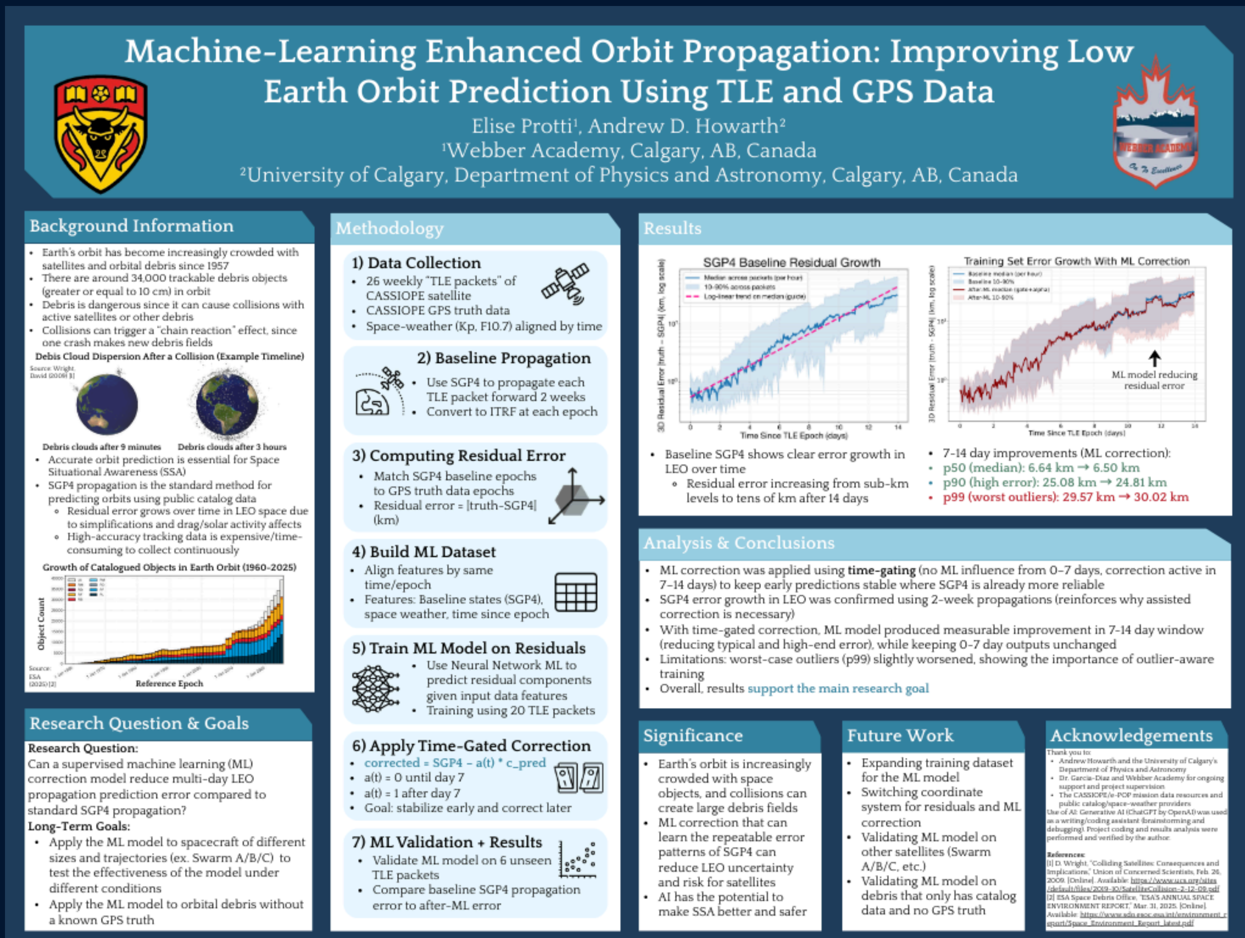
- Type 2 Diabetes: chronic high blood sugar -> DKD (Diabetic Kidney Disease) is a common complication which is the leading cause of kidney failure worldwide
- Animal models allow for a controlled/fair experiment to test for diabetic kidney disease
  - GK rats are preferred: spontaneously generate the same kidney damage that is seen in humans
- Goals:
  - prepare stained kidney tissue sections from healthy + diabetic GK rats
  - Capture high-res images of stained tissue
  - Perform ImageJ analysis
- Independent Variables:
  - Healthy controlled GK rates
  - Type 2 diabetic female GK rats
- Dependent Variables:
  - collagen content, glomerular cross-sectional area
- Significance:
  - Most studies use male rats instead of females -> want to find results for a more personalized treatment for females

Question(s):

- How do the GK rats get Type 2 diabetes?
- What is your sample size -> how many tests will you run?
  - Only going to have 1 trial, with 1 diabetic rat and 1 controlled rat and the left and right side of each kidney of each rat will be determined
- What is ImageJ software analysis?

🎤 Oral Presentation #2 - Science Fair Prep

FINALIZED POSTER DESIGN:



PLANNING:

- Suggested elements of the poster
  - Title and UofC/Webber crest
  - Background
  - Research question + long-term goals
  - Methodology
  - Results
  - Analysis
  - Conclusions
  - Significance
  - Future directions/work
  - Citations
  - References

Presentation #1: Andi

Notes:

- Automatic sarcasm detection -> systems trying to detect sarcasm in written English and speech
  - Limitations: computationally expensive, doesn't account for sarcasm types
  - Local features are understudied in sarcasm prosody (ex. elongation)
- RQ: Can a simple logistic regression model (using neural networks) detect sarcasm from speech?
- Used a MUStARD++ dataset (marked prominent words and sarcasm -> sarcasm type as well)
  - Measured Legendre coefficients and K-means clustering
- Experimented with lots of different types of models
- Results: no significance in any models for Legendre coefficients
  - Would be complicated to create a model that fits everything with significance
  - However, speech rate of prominent words showed a difference between sarcastic and sincere utterances

Question(s):

- What are the real-world applications of having an accurate sarcasm detection model?
- If you were to work with text based data instead of audio, how would your approach differ?
- Why did you not use Wavelet to automate the process?

Presentation #2: Ronald

Notes:

- Honeybees are a large industry in Canada -> Varroa mite parasites can cause bee colonies to die (very negative)
- Beekeepers must monitor the levels of mites -> beekeepers can look at the how many mites fell onto a sticky board to measure the degree/level of infestation
  - Use ML to count the number of mites on the sticky board

    - Want to maximize detection accuracy using only 64 labelled images (need to differentiate debris with the number of mites)
- Used a convolutional neural network
- Compare the neural network's outputs (boxes where it thinks mites are) to the actual area where a mite was in the image
- Relatively balances between false positives and false negatives
- Project helps to fill a geographic datagap in Canada

Question(s):

- How would you apply your findings/methods of improving accuracy of the ML model using scarce data to other fields/projects?

Presentation #3: Arvind

Notes:

- Cardiovascular physiology of stretching remains scarcely studied despite being so common
    - Stretch maneuver may lead to lightheadedness and syncopal induction
    - Significant hypotension in subjects
- Research questions:
    1) Does stretch induce transient, significant hypotension in tested subjects? -> YES
    2) Does prolonged stretch increase the magnitude of hypotension? -> YES
    3) Does orthostatic stress increase the magnitude of hypotension? -> YES
- Testing protocol: 2x supine and 2x standing stretch maneuvers for 30 seconds (neck twist , shoulder abduction, back hyperextension)
- SPSS software was used to measure the statistical significance
- Results -> hypotension IS induced by stretch -> is related to lightheadedness and syncope, orthostatic stress DOES impact the magnitude of hypotension
    - Stretch may be a significant factor in lightheadedness and syncope

Question(s):

- Do you think that the accuracy of your results would change if you had a larger sample size or were your results accurate enough with a smaller sample size (14-17 patients)?

Presentation #4: Shaayaan

Notes:

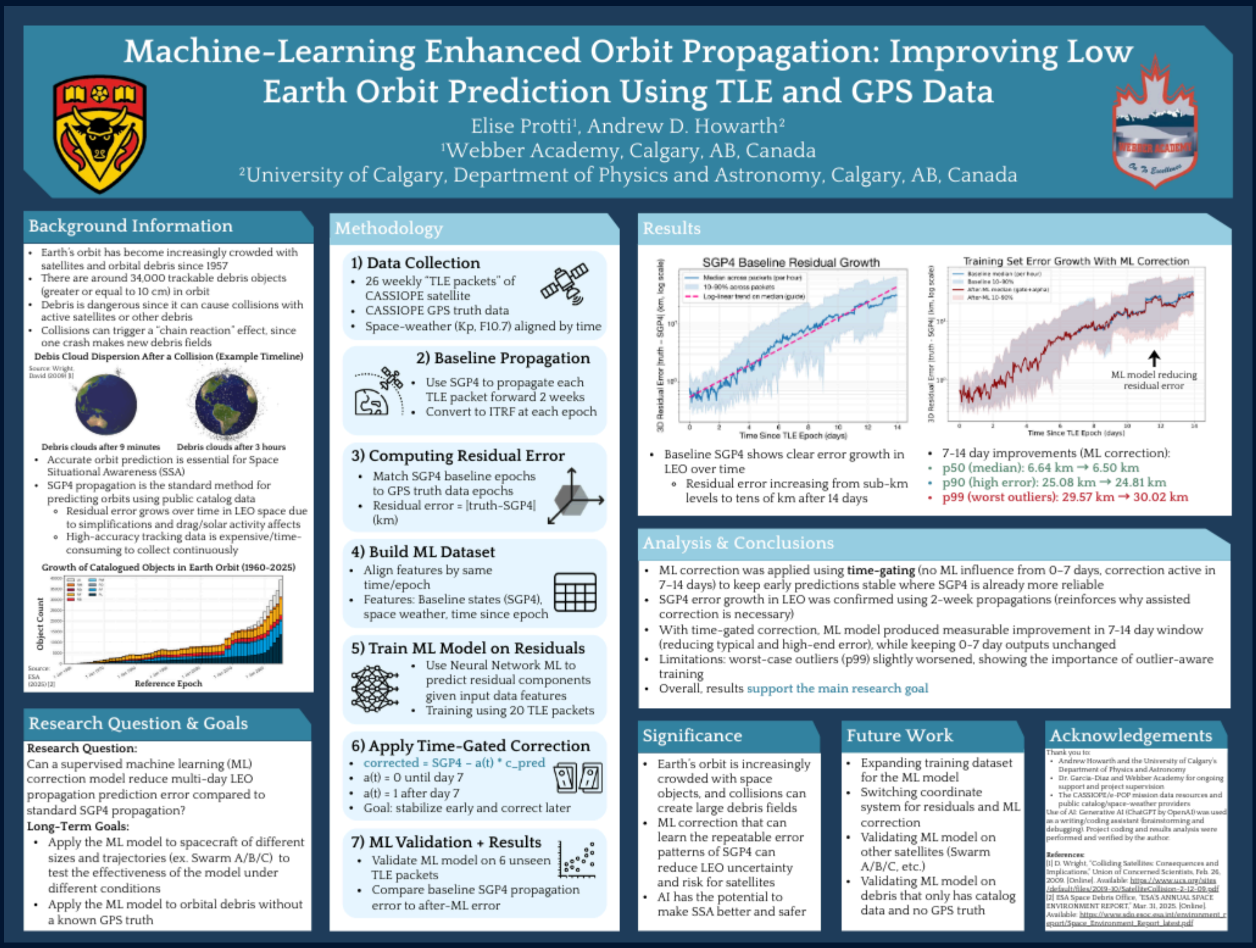
- He has 4 different samples of Al-SiC at different compositions to see how the microstructure changes and what properties changes
- Flexible matrix metal and super-strong ceramic reinforcement
- Need a lightweight and nonrusting sample for aerospace applications
- Created pellets to mimic the different properties of the alloys
- Hypothesis: as %W/W increases of SiC, the micro-hardness of the composite will increase proportionally -> act as shield for dust and debris
- Methodology steps: power to mixture (mortar and pestle for uniform dispersion of the powders) -> cold pellet press (turns powder to solid form) -> testing and analysis (microscopy for the 3D topographical map, nano indentation to test the deformation resistance)
  - AL-15SiC is resisting the diamond tip going further into the alloy -> this is the strongest (proved the hypothesis correct)

Question(s):

- How did you decide the percent compositions of the alloys that you used? Why didn't you use a 100% SiC substance?

📈 Science Fair Poster

SCIENCE FAIR POSTER



Elements of poster:

- Title and UofC/Webber crest
- Background
- Research question
- Methodology
- Results + Analysis
- Conclusions

- Significance
- Future directions
- Citations
- Acknowledgements

🔍 Results - Data Collection

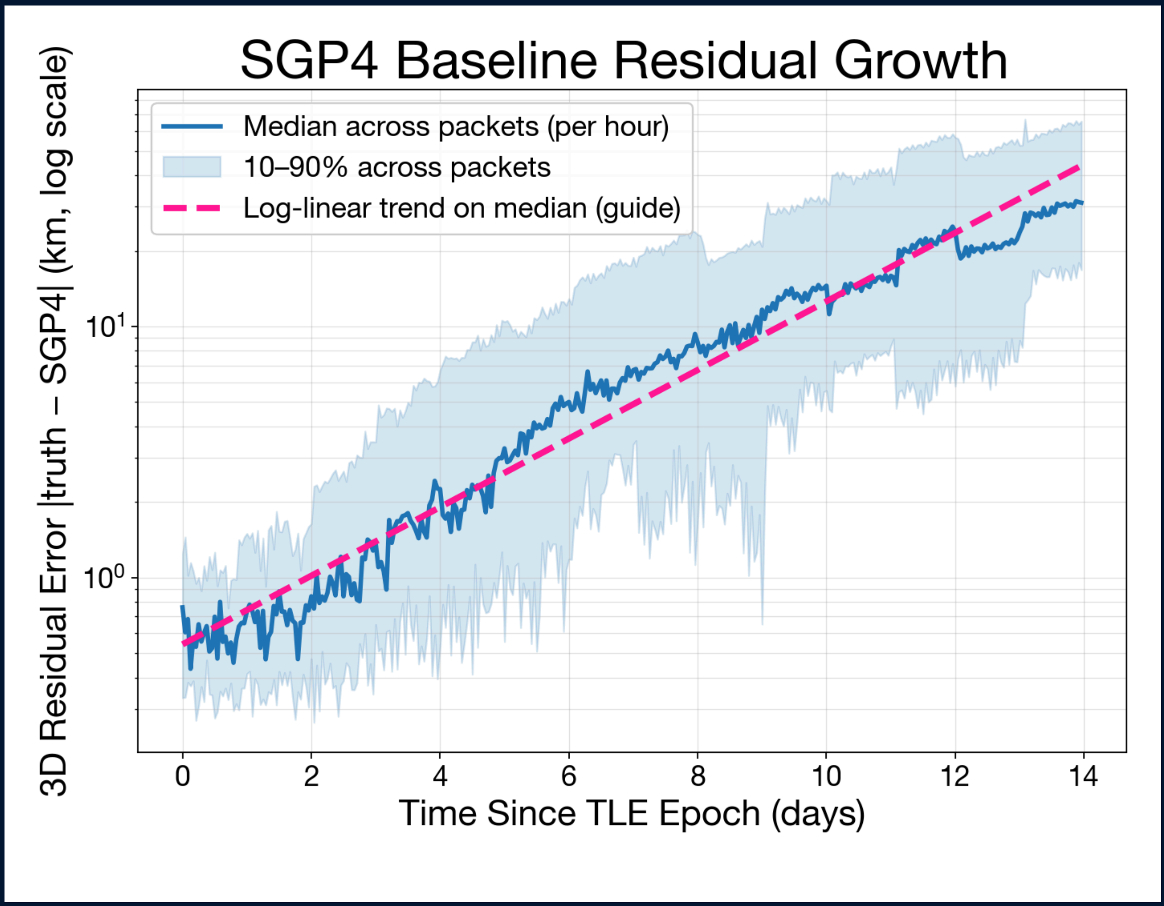
RESULTS - DATA COLLECTION

- ML-Correction Indicators from 7-14 days after epoch:
    - **p50 (median): 6.64 km → 6.50 km**
        - REDUCED ERROR FROM STANDARD SGP4 PROPAGATION
    - **p90 (high error): 25.08 km → 24.81 km**
        - REDUCED ERROR FROM STANDARD SGP4 PROPAGATION
    - **p99 (worst outliers): 29.57 km → 30.02 km**
        - INCREASED ERROR FROM STANDARD SGP4 PROPAGATION

- Graphical Analysis:

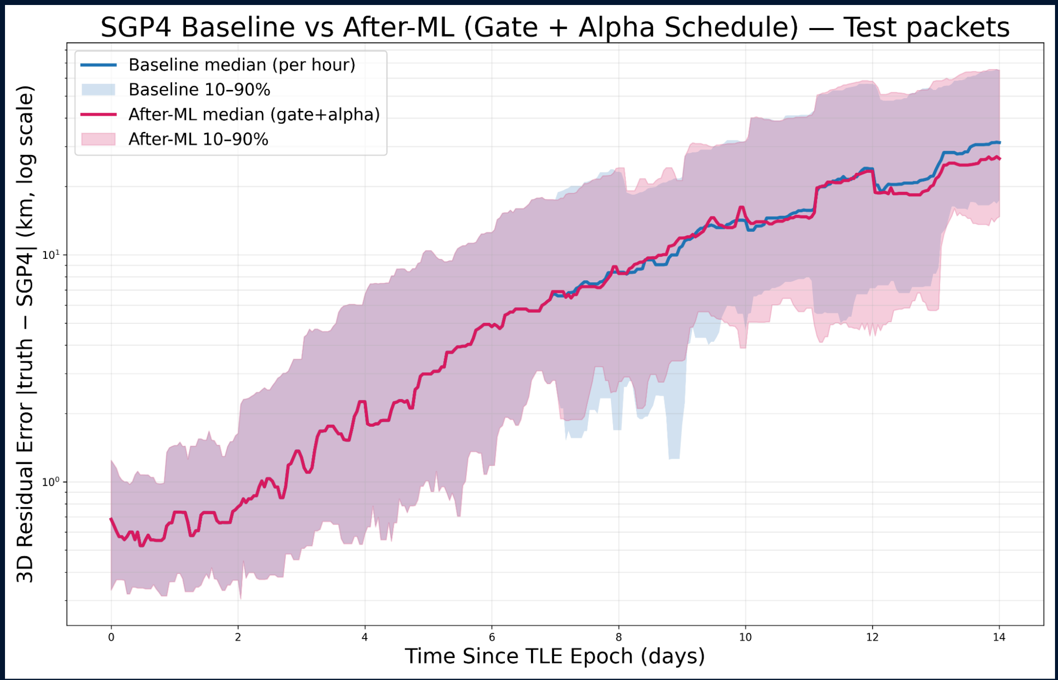
(1)SGP4 Baseline Residual Growth

- This graph shows the baseline SGP4 residual error growth when a TLE is propagated forward for 2 weeks (14 days) using the SGP4 propagation system. As the time since epoch increases, the 3D residual error resulting from the propagations increases exponentially, which is consistent with everything in scientific literature about SGP4 propagation error growth over time. This baseline SGP4 error growth serves as a baseline in all later graphs to compare the ML-corrections to (ideally the ML-corrections should have a lower residual error than the baseline SGP4 error level).



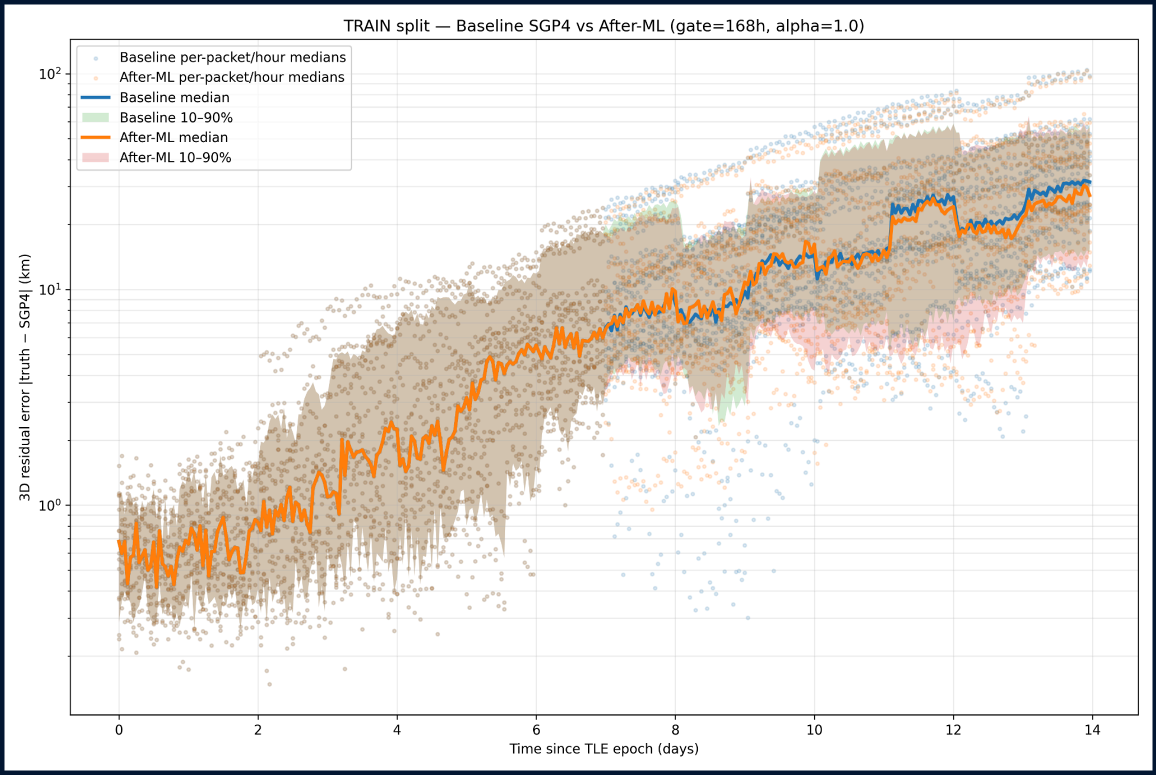
(2)SGP4 Baseline Error Growth Vs. After-ML Correction Error Growth (Test Packets) with Time-Gating from 7-14 Days

- This graph features percentile shading smoothing to look at the average pattern of error growth and where the ML model works best (most accurate predictions).



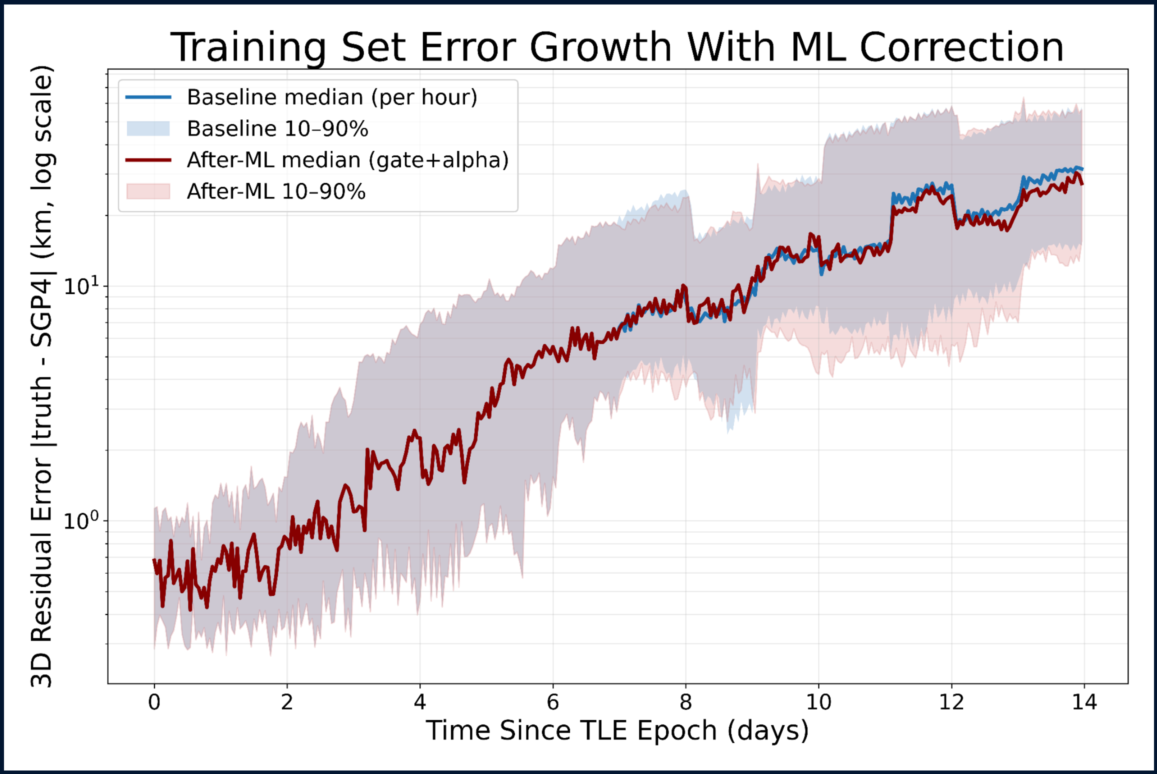
(3)SGP4 Baseline Residual Error Growth Vs. After-ML Correction Error Growth (Test packets) with Time-Gating from 7-14 days

- This graph features baseline and after-ML per-TLE-packet/hour medians shown by the coloured dots on the graph as well as the percentile shading (this helps to visually show the distribution of the data that the ML model outputted in comparison to the baseline SGP4 propagation values).



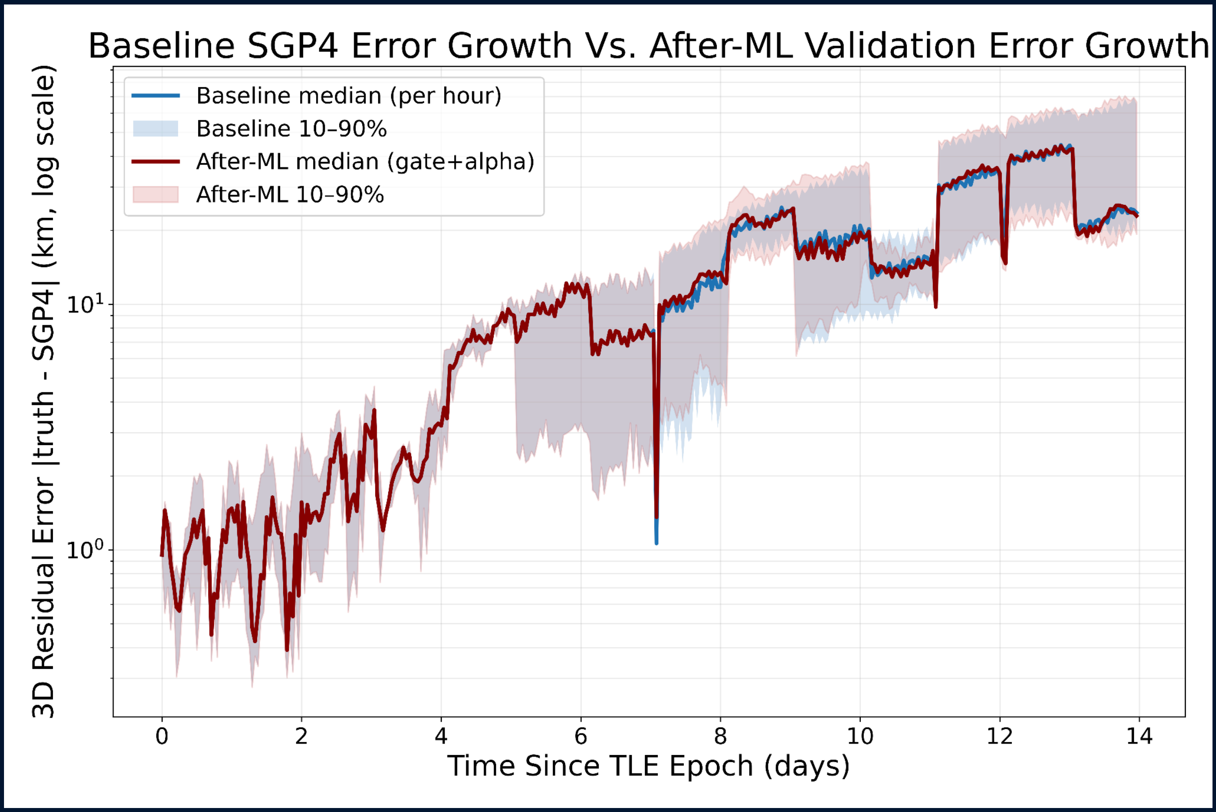
(4)Training Set Error Growth With ML Correction (ML Time-Gating active 7-14 days after initial epoch)

- This graph compares to error growth of the traditional baseline SGP4 propagations compared to the residual error growth after the ML model makes its corrections. It features lines for averages as well as percentage shading. In the graph, the ML correction line (amount of error) is clearly under the baseline SGP4 propagation average error line, showing the ML model was successful in reducing error arising from propagating the orbit/trajectory of the CASSIOPE satellite)



(5)Validation Set Error Growth With ML Correction (ML Time-Gating active 7-14 days after initial epoch)

- This graph compares the baseline SPG4 propagation error growth to the error growth with ML-correction on the validation/testing packets. Since the ML-correction and SGP4 baseline average error lines are approximately equal at all points in time after the 7 day mark (when time-gating is active), this could point to the conclusion that the model may be “overfitted” to the training data set, and will need to be trained on a larger, more diverse data set to improve prediction accuracy compared to the baseline SGP4 propagations level.



Baseline SGP4 Error Growth Vs. After-ML Validation Error Growth
Baseline median (per hour)
Baseline 10–90%
After-ML median (gate+alpha)
After-ML 10–90%
3D Residual Error |truth - SGP4| (km, log scale)
10^1
10^0
0
2
4
6
8
10
12
14
Time Since TLE Epoch (days)

💬 Analysis, Conclusions, & Significance

ANALYSIS, CONCLUSIONS, & SIGNIFICANCE

Analysis, Conclusions, and Significance sections from my Science Fair poster:

- More analysis will be added during the month of March when I am able to expand my project's focus to some of the FUTURE WORK I aim to achieve (expanding dataset for ML model, switching coordinate systems for residual error calculations, validating ML model on other satellites like Swarm A/B/C, validating ML model on debris with only catalog data and no known GPS truth).

Analysis & Conclusions

- ML correction was applied using **time-gating** (no ML influence from 0–7 days, correction active in 7–14 days) to keep early predictions stable where SGP4 is already more reliable
- SGP4 error growth in LEO was confirmed using 2-week propagations (reinforces why assisted correction is necessary)
- With time-gated correction, ML model produced measurable improvement in 7-14 day window (reducing typical and high-end error), while keeping 0-7 day outputs unchanged
- Limitations: worst-case outliers (p99) slightly worsened, showing the importance of outlier-aware training
- Overall, results support the main research goal

Significance

- Earth's orbit is increasingly crowded with space objects, and collisions can create large debris fields
- ML correction that can learn the repeatable error patterns of SGP4 can reduce LEO uncertainty and risk for satellites
- AI has the potential to make SSA better and safer