NASA Conjunction Assessment Risk Analysis (CARA) Compendium for Artificial Intelligence and Machine Learning for Satellite Collision Avoidance

Alinda K. Mashiku
NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771
Lauri K. Newman
NASA Headquarters, 300 Hidden Figures Way SW, Washington, DC 20546

Dolan E. Highsmith
The Aerospace Corporation, 14745 Lee Rd, Chantilly, VA 20151

ABSTRACT

To ensure the continued and sustainable use of Low Earth orbit (LEO) and beyond, reliable and effective collision prediction between space objects is critical as the orbital environment faces unprecedented growth. Space Situational Awareness (SSA) for commercial and government missions now includes rapidly expanding satellite populations ranging from small, potentially less agile CubeSats to large constellations, each presenting unique conjunction assessment challenges and collision risks in the orbital regimes. With space object catalogs expected to increase tremendously and conjunction events becoming more frequent and complex, traditional risk assessment methods may face significant scalability limitations. NASA's Conjunction Assessment Risk Analysis (CARA) program has conducted comprehensive research utilizing over 450,000 Conjunction Data Messages (CDM) from 2015-2018 to evaluate whether artificial intelligence and machine learning capabilities could enhance or augment collision avoidance decision making by performing conjunction assessment with increased speed or certainty using CDM data available early in event timelines. In essence, can these approaches enable rapid, reliable decision-making for collision avoidance within the critical 7-day identification window, incorporating expected or predicted information from anticipated new observations that could provide a bounded solution space? The research employed multiple Artificial Intelligence/Machine Learning (AI/ML) methodologies including supervised learning, unsupervised clustering techniques, Fuzzy Inference Systems, Deep Neural Networks, and Long Short-Term Memory models to identify risk-associated patterns during a conjunction as well as to analyze the temporal evolution of defined risk parameters. Statistical and information theory parameters were investigated beyond traditional probability of collision calculations, developing adaptive models for varying CDM availability, and implementing sophisticated time-series analysis approaches. The study identified fundamental challenges with the operational use of AI/ML for conjunction risk assessment, including data scarcity, the stochastic nature of orbital mechanics, model interpretability requirements, and the critical need for explainable AI approaches that can meet the high-reliability standards essential for space operations decision-making.

1. INTRODUCTION AND PROBLEM STATEMENT

Performing Space Situational Awareness (SSA) for commercial and government missions is now complicated by the rapidly growing number of operating satellites of varying capabilities, ranging from small and potentially less agile satellites to multitudes of large maneuverable satellites with diverse maneuverability capabilities. The space object catalog has increased tremendously in size, such that in the very near future, with ongoing projects having intentions of adding thousands of objects to the current orbital environment, traditional approaches for conjunction assessment and collision avoidance need to be revisited to meet the scale of demand. This exponential growth in the number of space objects has rendered the traditional "big sky theory" obsolete, creating an urgent need for state-of-the-art risk mitigation practices and early collision assessment capabilities for efficiency and cost savings through the use of precise data.

The NASA Conjunction Assessment Risk Analysis (CARA) program has served as the primary authority for non-human spaceflight missions at NASA in defining metrics for safe and unsafe conjunction assessment design and process approaches, as well as in providing conjunction assessment and risk mitigation analysis and support to spacecraft

operators [15]. The CARA program relies on input data such as CDMs and uses statistical parameters such as probability of collision (Pc) calculations to evaluate conjunction risks, with the level of risk vacillating with the time-varying and stochastic nature of state uncertainties as space objects are propagated and new observations become available. CARA performs conjunction assessment from screened spacecraft trajectories that are predicted as far as 7 days out, to assess potential collision risks. However, critical decisions to implement collision avoidance maneuvers typically occur within a couple of days to several hours prior to the Time of Closest Approach ((TCA), when the two objects are at their closest distance separation)), creating significant time pressure for rapid and accurate risk assessment.

The most accurate method for performing conjunction assessment between two spacecraft involves using Monte Carlo runs of the state space of both the primary and secondary objects. However, even when employing parallel computing and high-accuracy prediction models, Monte Carlo runs are not feasible for routine screenings of the space catalog because of their prohibitively long computation times and computational burden for analyzing the full catalog. However, Monte Carlo runs can be used to analyze specific conjunction events to estimate collision and conjunction rates for a particular object or conjunction event [15]. As a result, the 2-D probability of collision calculation is a widely accepted metric for risk assessment. This approach projects the close approah into the conjunction plane, with the HBR circle at one end of the estimated miss vector and the combined uncertainty at the other end; and it calculates the likelihood that the miss distance, when perturbed by a random draw from the combined uncertainty, will actually terminate within the HBR circle (see Fig 1). The conjunction plane is the plane that is normal to the relative velocity vector of the primary and secondary objects. While these computations can be performed reasonably quickly under Gaussian assumptions, the associated two-dimensional assumptions may fail to hold in the realm of low-relative-velocity encounters, where objects move more slowly through each other's uncertainty volume.
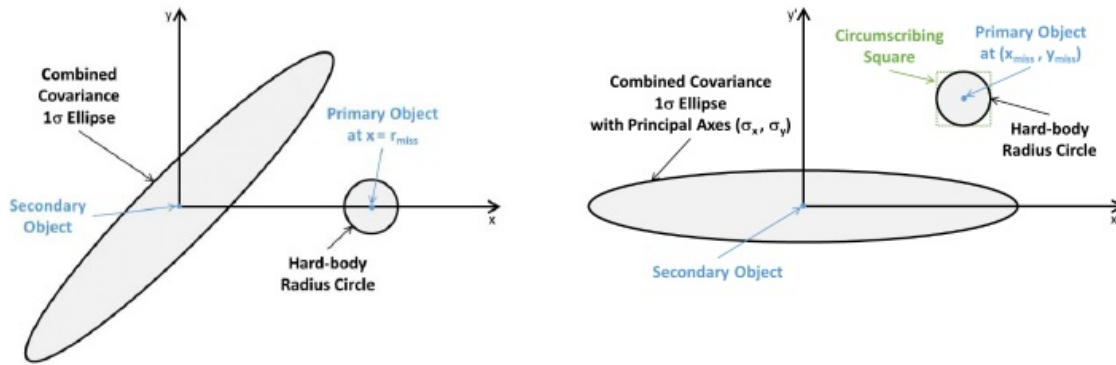

Fig. 1: 2-D Conjunction Plane showing total covariance and total Hard Body Radius Ref.[15]

The limitations of current methods become particularly pronounced when dealing with inaccurate covariances, which can lead to misguided risk margins for events and therefore result in the prioritization of non-close approach events over legitimate risks [1, 2, 7]. Moreover, the assumption of Gaussian-error ellipsoids can be problematic, as nonlinear dynamics significantly affect spacecraft trajectories. Current models for miss-distance prediction provide error bounds, but these bounds may not encapsulate the entire risk spectrum and may also introduce high false alarm rates [4]. With the expected dramatic increase in space objects, the complexity of managing large constellations, and the growing number of CDMs, it became pertinent to investigate the potential for using AI/ML capabilities for SSA and CA.

Therefore, NASA CARA initiated a research endeavor to determine whether artificial intelligence and machine learning capabilities could be used to aid decision makers in making faster and more accurate assessments, or to determine any physics-based patterns for close conjunctions using AI/ML techniques on the large available dataset of close approaches. The motivation for this investigation stems from the potential for machine learning algorithms to enhance collision risk parameter association and provide early information for decision-making through both unsupervised and supervised learning approaches. This research represents a critical step toward identifying the required next-generation conjunction assessment capabilities and data quality required, that would be needed to effectively handle the increasing complexity and volume of space traffic while maintaining the high reliability standards required for space operations.

2. BACKGROUND AND CURRENT STATE OF AI/ML FOR SPACE SITUATIONAL AWARENESS

Between inception in January 2005 and June 2024, CARA has performed routine CA screenings resulting in more than 11 million CDMs, creating a substantial historical dataset of predicted close approaches of varying risk levels that could potentially benefit from advanced analytics techniques. CDMs serve as the primary data source for space object collision risk assessment, containing points of information about the relationship between primary and secondary objects during a conjunction, including state vectors, probability of collision (Pc), time of closest approach (TCA) and other related parameters. These data messages are typically available three times a day, with the earliest possible data collection typically occurring about seven days before TCA. This large quantity of available data has driven an interest in applying AI/Ml to the CA problem. However, most high-risk conjunction events on average have fewer than 7 CDMs available throughout the entire event timeline. This data scarcity becomes even more pronounced when examining close approaches within a 5km radial miss-distance, resulting in a subset of events now having less than 4 CDMs per conjunction event. Therefore, although there is a large quantity of available data, the data required for high-risk conjunction assessments for each event can be limited, and given that the CDMs and each conjunction is unique, this situation poses challenges for use of AI/ML.

CDMs are generated no earlier than 7-days prior to TCA for low Earth orbit due to inherent limitations in orbit prediction accuracy when propagated over longer periods of time, predominantly due to atmospheric drag effects on the trajectory causing position and velocity uncertainties. Any data generated from longer propagation may be inaccurate for reliable decision-making. Additionally, conjunction identification within 8 hours of TCA (called "late notice events") can and does occur due to unpredicted changes in atmospheric drag caused by solar storms or other space weather phenomena. This limited time horizon creates significant challenges for risk assessment and collision avoidance decision-making, as operators often have minimal data points upon which to base critical maneuver decisions. The combination of few available CDMs per event and the short timeline for analysis constrains the development of reliable prediction models, which pose challenges for utilizing AI and ML approaches for conjunction assessment and decision support.

Current collision probability computation methods are heavily based on statistical approaches that make significant assumptions about the geometry of spacecraft encounters. Figure 2, shows a sample Case Study of three distinct satellite conjunction scenarios observed during a period of significant solar activity, demonstrating the dynamic nature of and collision risk assessment due to uncertainties in space object tracking and space weather prediction.
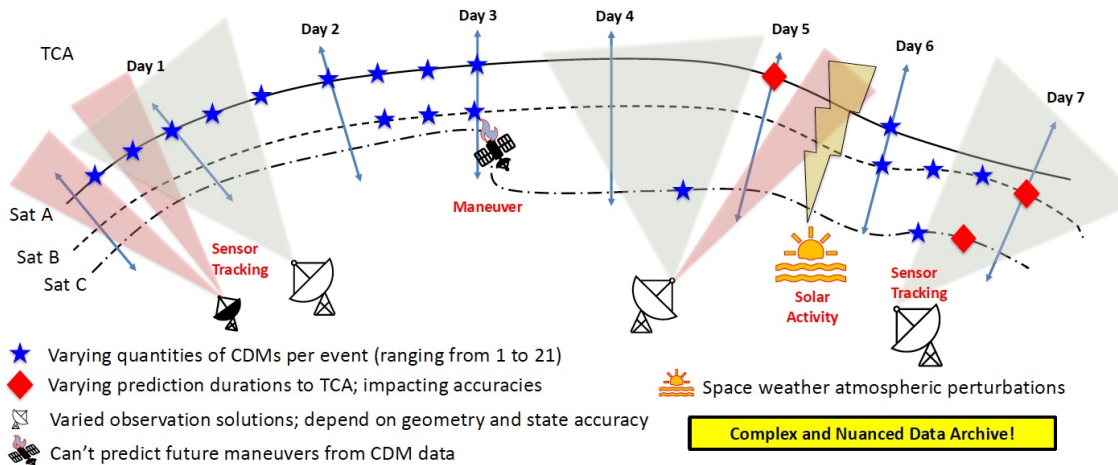

Fig. 2: A visual capture demonstrating multiple evolving factors in conjunction assessment timeline [13]
.

2.1 Satellite A - Conjunction Profile:

Conjunction screening indicated a high-risk close approach for Satellite A on Day 5 beginning on Day 1, and despite ground-based sensor networks collecting updated tracking data at each point identified with a blue star icon, the close approach continued to be predicted through Day 3, when updated orbit determination around Day 4 led to the

conjunction no longer being a threat. However, following the passage of the predicted TCA on Day 5, solar activity disrupted the object and hence the tracking updated values, causing the reappearance of a conjunction with the identical object on Day 6. The conjunction risk ultimately resolved without intervention.

2.2 Satellite B - Risk Evolution:

Satellite B first received conjunction notification on Day 2 of the monitoring period. The identified risk appeared to resolve naturally by Day 3. However, following the solar activity event on Day 5, the conjunction reappeared from the tracking systems with critically reduced response time—of only one day remaining for collision avoidance planning prior to the revised TCA on Day 7.

2.3 Satellite C - Maneuverable Asset Response:

Satellite C planned to execute a routine maneuver (i.e. station-keeping, Drag Make Up (DMU) etc.) on Day 3. Immediately following the completed maneuver, the spacecraft encountered a single conjunction event on Day 4 that, after analysis, was concluded to be at the edge of the safety volume used to compute close approaches and not a significant high-risk event. By Day 5, it had resolved itself. However, after the solar activity event subsequently caused a high-risk conjunction event with significant minimal advance warning — only several hours prior to the calculated TCA.

These cases illustrate the uniqueness of each satellite conjunction event, the significant impact of space weather events on conjunction assessment accuracy, the varied instances of conjunction notifications, and the overall challenges posed to satellite operators in maintaining situational awareness during periods of enhanced solar activity. The complexity also demonstrates why it would be desirable to determine whether AI/ML techniques can identify novel computational approaches through physics-based pattern detection in close-conjunction scenarios from the large CDM datasets of historical high-risk conjunction events to enable decision-makers to achieve faster conjunction risk level determinations, if the predicted perturbations can be quantified. They also demonstrate the critical importance of establishing explainability frameworks for any AI/ML model outputs (also known as Explainable AI) intended for operational collision avoidance applications so that decision makers can understand and use the findings as part of their process.

Training AI/ML systems to augment Conjunction Assessment Risk Analysis (CARA) operations necessitates access to verified ground truth data. While comprehensive collision datasets remain limited, the theoretical assessment of close conjunction events characterized by non-zero miss distances provides a viable foundation for evaluating the operational validity of AI/ML implementation in conjunction assessment workflows.
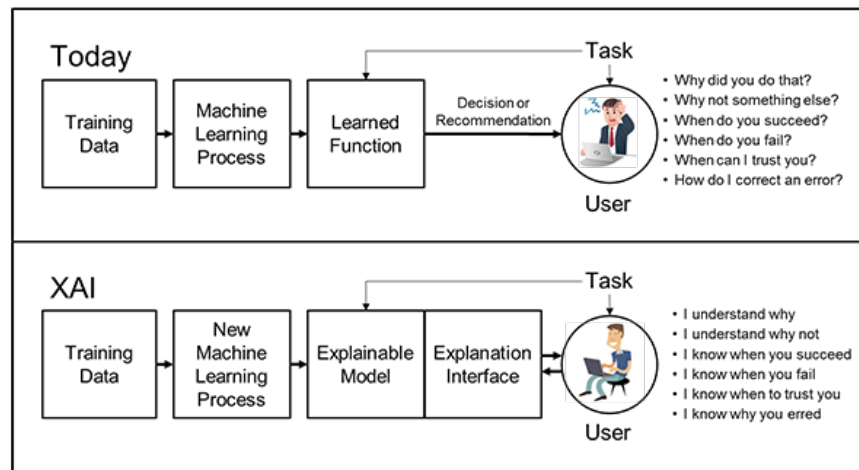

Fig. 3: A visual capture demonstrating multiple evolving factors in a sample conjunction assessment timeline [11]
.

Explainable AI implementation requires operators to comprehend both the underlying model architecture and the interpretation of interfaces and outputs generated by AI/ML systems, particularly in high-risk, high-value operational environments. Conjunction event assessment involves multiple non-deterministic variables, and given the inherently

unique characteristics of each conjunction scenario, the development of AI/ML-powered decision support tools mandates the incorporation of robust explainability frameworks. These frameworks must provide transparent insights into model reasoning processes, enabling operators to validate AI/ML recommendations against established operational protocols and maintain confidence in automated decision support capabilities during critical collision avoidance operations, see Fig 3.

3. MACHINE LEARNING APPROACHES INVESTIGATED

Machine learning methodologies can be broadly categorized into supervised, unsupervised, and reinforcement learning paradigms, each offering distinct perspectives across diverse fields and applications with potential for methods or approaches for performing conjunction assessment analysis and information parameters in decision-support for risk mitigation. Supervised learning methods utilize labeled training data to learn relationships between input parameters and desired outputs, such as, in our scenario, classifying conjunction events as either "safe" or "unsafe" or the ability to predict specific risk metrics like miss distance. Unsupervised learning techniques analyze data without predetermined labels to discover hidden patterns, clusters, or relationships within the conjunction data that may not be immediately apparent through traditional analysis methods.

Reinforcement learning approaches were not extensively explored in these investigations. The focus remained primarily on supervised and unsupervised methods due to their natural potential applicability to the conjunction assessment problem, which can be framed as both a classification problem (determining safe or risk levels through risk-associated parameter combinations) and a pattern recognition problem (identifying specific patterns in a predisposed high-risk conjunction event).

The investigation of statistical and information theory parameters represented a potentially insightful perspective as counterpoint to the classical probability of collision computations, to determine the feasbility of incorporating a broader set of risk-relevant metrics. Seven key information parameters were implemented, including probability of collision, miss distance, Mahalanobis distance, Bhattacharyya distance, and Kullback-Leibler distance, among others, to capture and study the different aspects of conjunction geometry and close-approach information measures. These parameters were designed to supplement traditional Pc calculations by providing additional statistical insights into the nature of each encounter as described in detail in [14]. The approach aimed to generate augmented statistical parameters that could characterize the collision risk by incorporating information theoretical parameters and statistical measures for an early decision-support paradigm.

3.1 Unsupervised Machine Learning

Unsupervised machine learning clustering techniques were implemented to identify natural groupings within conjunction data and discover parameter associations that could inform close approach risk inference. The unsupervised approaches offered the advantage of discovering hidden patterns without requiring pre-labeled training data. The intention of this initial approach of using unsupervised machine learning on the information parameter set was to explore the data and determine whether an internal representation existed that would cluster the parameter sets into two categories: safe or close encounter (not safe). Methods such as K-means clustering and Support Vector Machines were investigated in this effort [14].

The clustering method also considered in this work associated a series of CDMs in a time-series via the agglomerative hierarchical clustering approach. This method is generally successful in performing high-dimensionality reduction and feature extraction in other applications [5, 18]. Due to the number of dimensions assessed for risk assessment using the data available in the CDM and other space weather data, hierarchical clustering was selected to identify clusters to reduce the number of dimensions of the parameters down to the most pertinent ones required for conjunction assessment. The hierarchical clustering implementation works from a "bottom-up" approach, where each data point is associated with its neighbor with the closest norm distance [18], where at each level of a dendrogram, a horizontal slice reveals the number of clusters and points in a given cluster[16]. The results of this operation are displayed in a dendrogram, which is a tree-like structure as seen in Figure 4.

Competitive networks were implemented to initialize the weight vectors that trained iteratively on input data to identify clusters and parameter associations for the hierarchical clusters. To validate the trends identified in the hierarchical clustering, the competitive network models were initialized with a decreasing number of weight vectors, such that strongly-related data still attracted a weight vector. Pre-processing data, such as standardization, ensures that the
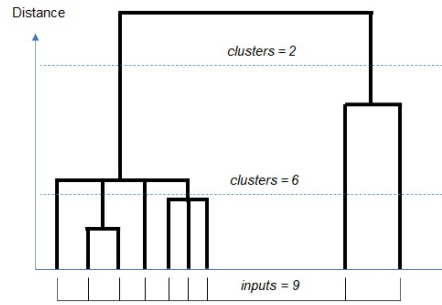
Fig. 4: Hierarchical clustering results displayed in a dendogram. In this example, nine inputs are clustered. Horizontal slicing for the number of clusters = 2, 6 cases are indicated.

.

network is learning the data rather than the numerical values so that parameters that are not important or misleading in a data set are identified and appropriately ignored [16]. These clustering methods were chosen to reduce the dimensionality of the conjunction assessment problem and identify which parameters naturally separated events into high-risk categories.

3.2 Fuzzy Inference Systems

Fuzzy Inference Systems (FIS) were investigated as a method to process multiple risk parameters through weighted assignments with pre-defined thresholds, addressing the inherent ambiguity in conjunction assessment decisions. The FIS model was chosen because it could capture partial memberships of variables into different risk categories and generate continuous outputs rather than discrete classifications, which was particularly insightful for highlighting decision ambiguities that exist when risk parameters fall within close neighborhoods of established thresholds. These systems can be trained or untrained and constructed manually or automatically through learning on a dataset. This study focused mainly on manually constructed systems that were trained after unsupervised classification for model fitting purposes. Figure 5 below shows a snapshot of the graphical user interfaces (GUI) of the FIS using MATLAB's Fuzzy Logic Toolbox [10]. The system generated weights based on defined thresholds of each parameter to assess conjunction event outcomes, incorporating results from unsupervised machine learning clustering techniques using K-means and Support Vector Machine methods to determine parameter correlations and obtain appropriate weights. The detailed description and approach is provided in [14].
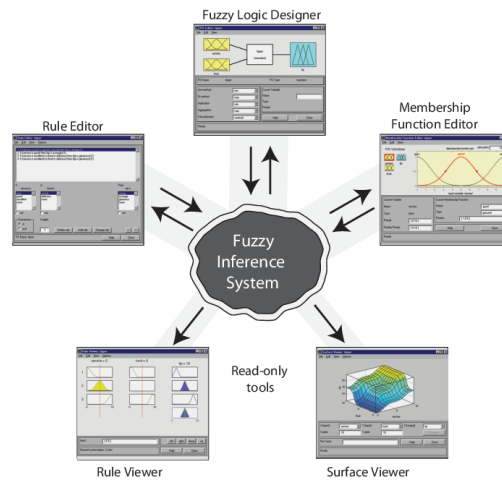

Fig. 5: Fuzzy Inference System User Interface using MATLAB[10]

.

3.3 Deep Neural Networks

Deep Neural Networks (DNN) refers to a neural network architecture that can have several hundreds of hidden layers implemented in the model (see Fig 6), thus automatically implying an unlimited accuracy potential [9]. However, this implies the availability of sets of data that are both enormous in size and precise in information capture. The DNN approach has demonstrated success in pattern recognition and its ability to learn intricate relationships from large datasets without requiring explicit feature engineering in various applications. DNNs were employed to characterize the evolving nature of CDMs due to the stochastic characteristics of conjunction events, with the goal of capturing complex, non-linear relationships within the conjunction data that traditional methods might miss. The details of the informational parameters, groupings, and back-propagation algorithms considered are detailed in the following reference [14].
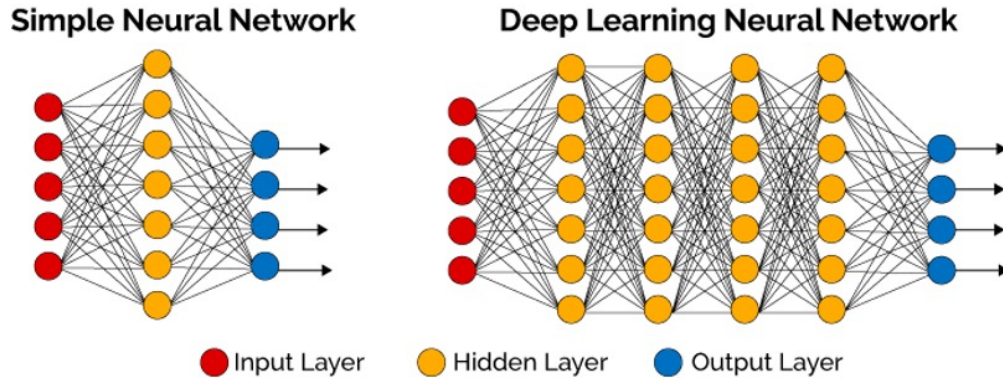

Fig. 6: A simple neural network compared to a deep learning neural network based on the numbers of hidden layers [20]

.

3.4 Supervised Machine Learning Approaches with Time-Series Capture (Regression Neural Networks [RNNs])

Given that conjunction data for an event timeline with multiple CDMs can be interpreted and analyzed in a time-series, the adoption of using supervised learning enables us to focus on connecting features or the information parameters to decision outcomes and thus, theoretically, aide in predicting the future state or risk level of a conjunction. Shallow neural networks are a type of neural networks that use simple mathematical equations to find which inputs best predict a specific outcome, but they operate at a linear level of complexity and thus can be susceptible to errors from highly nonlinear coupling. In addition, standard neural networks (either shallow or deep), do not remember or account for previous time-series inputs but only focus on the immediate past data during the training phase. Therefore, setting the premise for a memory-based network that can incorporate the historical available data to aid in prediction with an expected performance, accuracy, and more importantly, the potential for added explainability. Regression Neural Networks (RNNs) are a type of neural network that retains memory about an input and utilizes that associated memory value to predict the outcome on a time-dependent series [8, 17]. The Long Short-Term Memory (LSTM) network is a type of RNN that was implemented to handle the time-varying nature of CDM data as the conjunction assessments evolves toward TCA.

The main difference between an RNN and an LSTM, see Figure 7, is the length of time that information can be maintained in memory [19]. The set of gates in the LSTM Unit is used to control information within memory to regulate what information should enter (Input Modulation Gate), how long it should be kept, what the decay rates in the Forget Gate should be to start ignoring stale information to ultimately allow the model to provide the output [19].

The LSTM architecture was particularly appealing because it could retain memory about each CDM data set in an event and potentially learn from the temporal evolution of the conjunction parameters. Specifically, the LSTMs selectively retain or discard information, ensuring the network can focus on relevant features and eliminate irrelevant data. Given that CDM data for conjunction events have varying lengths, the LSTM can adapt to sequences of varying lengths, allowing them to be flexible and robust for handling diverse sequential data [16].
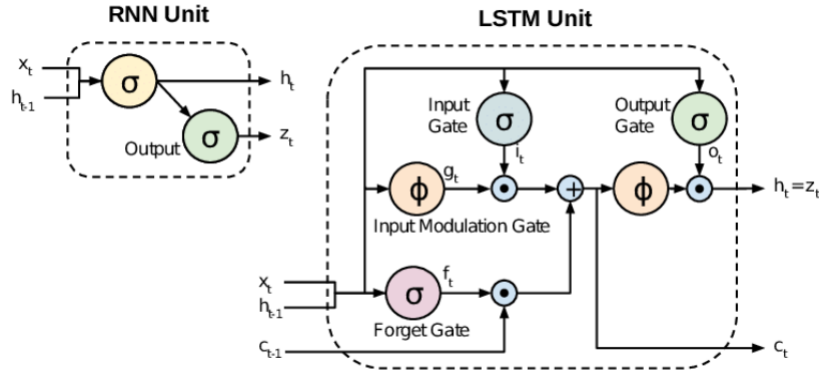
Fig. 7: A simple neural network compared to a deep learning neural network based on the numbers of hidden layers [19]

.

4. METHODOLOGY, DATA ANALYSIS AND RESULTS

4.1 Dataset Overview and Preparation

The research utilized a comprehensive dataset of approximately 450,000 historical conjunction cases from the NASA CARA program archives collected from 2015-2018, theoretically providing a robust foundation for statistical analysis and model training. To address the challenge of few high-risk events in the historical data, scaling factors were implemented that decreased the probability of collision (Pc) threshold used to identify high-risk events to include a larger number of events in the analysis along with their corresponding informational parameters such as miss distances, relative velocities, approach angles, and orbital eccentricity variations. The cases spanned multiple orbital regimes including Low Earth Orbit (LEO), Geosynchronous Earth Orbit (GEO), Medium Earth Orbit (MEO), and High Earth Orbit (HEO) to ensure comprehensive coverage of different conjunction scenarios. Additionally, in some analyses the dataset was augmented with 3,998 simulated test cases generated from eight baseline data cases from [1, 3], by randomly varying state and covariance information at TCA with 500 sample cases derived from each baseline case to increase statistical robustness [3].

The data preprocessing involved extensive normalization procedures to handle the varying orders of magnitudes across the different numerical parameters. All individual CDM parameters were normalized to a range of 0-10, with outliers systematically removed to prevent saturation of values at the extremes. The dataset incorporated 18 key statistical parameters and space weather data including miss distance, relative velocity, drag coefficient, solar radiation pressure coefficient, energy dissipation rate, probability of collision, hard body radius, time to TCA, approach angle, Mahalanobis distance, and space weather indices (F10.7, AP, DST). For events restricted to 5km radial distance, the dataset was further constrained, resulting in most events having fewer than 4 CDMs per event, which significantly impacted the viability of training data for machine learning models. Details on the information parameter generation are presented in detail in [14, 3, 16].

4.2 Data Analysis and Results

4.2.1 Unsupervised Machine Learning: Classification and Clustering Methods

This portion of the investigation employed both K-means clustering and Support Vector Machine (SVM) classification techniques to evaluate six information parameters derived from conjunction assessment scenarios (Probability of Collision, Miss-Distance, Mahalanobis Distance, Battacharya Distance, Kullback-Leibler Distance, and Orbit Angle) [14].

K-means clustering K-means clustering partitions data into clusters based on the closest mean of a chosen number of centroid points, which are representative of classes. For our K-means clustering, we chose 2 centroids (K=2)

for each parameter, representing a "close encounter/not safe" and a "safe" categorization. Execution of K-means clustering resulted in the following value outputs:

1. K-means Centroid Values corresponding to the classes

2. K-means Standard Deviations around the centroids

3. K-means Performance Metric (PM) generated by calculating the percentage of the dataset that was correctly assigned to the ground truth (see Table 1)

Support Vector Machines (SVM) clustering SVM-clustering techniques separates data into two classes using kernel methods to extend its dimensionality in order to find separations in the data that are not existent in its distribution from lower dimensional spaces [14]. Execution of SVM classification methods resulted in the following values:

1. SVM Standard Deviations around the centroids

2. SVM Performance Metric (PM) generated by calculating the percentage of the dataset that was correctly assigned to the ground truth (see Table 1)

Parameter	K-means (PM)	SVM (PM)
Probability of Collision (Pc)	0.7742	0.9995
Miss Distance (MD)	0.6389	0.8314
Mahalanobis Distance (MHD)	0.6983	0.8810
Battacharyya Distance (BD)	0.7611	0.8864
Kullback-Leibler Distance (KLD)	0.7736	0.8459
Orbit Angle (OA)	0.5387	0.8711

Table 1: Clustering Performance Metric (PM) for K-means and SVM Methods

Performance metrics were established by comparing algorithmic outputs against discretized ground truth classifications based on Monte Carlo simulation results. SVM classification demonstrated superior performance across all evaluated parameters, achieving correlation coefficients ranging from 0.8314 to 0.9995, with the probability of collision parameter yielding the highest correlation at 0.9995. In contrast, K-means clustering exhibited lower performance metrics, with correlation coefficients ranging from 0.5387 to 0.7742. The Mahalanobis distance, Bhattacharyya distance, and Kullback-Leibler divergence parameters showed promising correlation potential, while the orbit angle parameter demonstrated the lowest correlation with ground truth classifications. The superior performance of SVM methods led to their adoption for informing the construction of subsequent Fuzzy Inference System architectures in the next section.

4.2.2 Fuzzy Inference System (FIS) Model

Three types of FIS models were tested - Mamdani, Sugeno, and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) - to augment the satellite collision avoidance decision-making process in addition to the traditional Pc omputations. The Mamdani FIS model was manually constructed using fuzzy membership functions (FMF) that map input variables to output decisions through "if-then" rules with partial truth values between 0 and 1, rather than discrete Boolean logic. The Sugeno model followed similar inference processes but utilized linear or constant output membership functions, while ANFIS incorporated neural network-like training mechanisms to automatically infer optimal membership function structures [10].

In the prototype implementation, we constructed a Mamdani FIS using three information parameters derived from Alfano's simulated dataset [1]: miss distance, probability of collision, and Kullback-Leibler divergence (KLD). The system was designed with weights derived from Support Vector Machine clustering outputs from Section 4.2.1, creating a multi-parameter decision-making tool that outputs binary classifications (0 for unsafe, 1 for safe encounters).

Fig. 8: FIS Rule Viewer providing the decision output (Large Miss-Distance, Low Pc, provides a decision output of 0.837 (Decision = 1 for safe)[14]
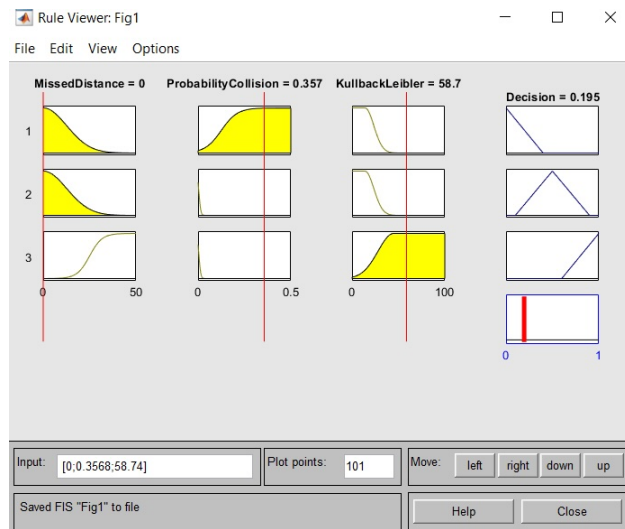
.

Fig. 9: FIS Rule Viewer providing the decision output (Low Miss-Distance, High Pc, provides a decision output of 0.195 (Decision = 0 for unsafe)[14]

.

However, the performance revealed significant limitations, particularly with the KLD parameter, which failed to contribute meaningful information to the final decision output despite having high SVM weight values, see Figures 8 and 9

Overall, the FIS approach demonstrated less than desired performance based on the parameter selection and limited practical applicability for collision avoidance decision support because it lacked the necessary reliably deterministic and explainable outcome that would be needed by decision-makers. We discovered that the KLD parameter, which compared probability density functions between primary and secondary objects, did not provide direct risk value despite its conjunction assessment statistical significance. This highlighted a critical finding that not all statistically weighted parameters contribute useful information for close approach prediction when applied universally across all conjunction events, and that careful physics-based parameter selection is essential. This FIS assessment concluded that more investigation is needed to determine optimal parameter combinations for a FIS application, as many informational parameter groupings failed to produce intuitive or expected outputs, leading us to investigate Deep Neural Network approaches, as will be described below.

4.2.3 Deep Neural Network

The Deep Neural Network (DNN) implementation utilized supervised learning approaches with four distinct parameter groupings (based on FIS performances) that were tested across architectures of 10, 20, and 40 hidden layers. Two backpropagation training algorithms were evaluated: the Levenberg-Marquardt (LM) algorithm, selected for its superior performance and computational efficiency, and the Scaled Conjugate Gradient (SCG) algorithm, chosen for its enhanced performance among gradient-based methods.

The LM algorithm was designed to reach second-order training speed without needing to compute the Hessian, which is a second-order derivative matrix, or the derivative of the Jacobian[12, 6, 21]. The LM algorithm uses the following approximation to the Hessian matrix **H** in a Newton-like update as shown in Equation 1 and 2 [6]:

$$w_{k+1} = w_k - [J^T J + \mu I]^{-1} J^T e \tag{1}$$
$$H = [J^T J + \mu I] \tag{2}$$

where J is the Jacobian matrix that contains the first derivatives of the network errors with respect to the weights **w** and biases. **e** is a vector of the neural network errors, I is the identity matrix and μ is known as the combination coefficient [6, 21]. During training, the scaled conjugate algorithm predefines several stopping conditions such as the number of epochs (iterations), maximum duration, performance goal, minimum performance gradient, and validation performance [14].

The networks employed standard training/validation/testing ratios of 70%, 15% and 15% respectively, with performance measured using regression R-values and Root Mean Square Error (RMSE) metrics on a dataset of 1,000 simulated samples containing both safe and close encounter classifications [1, 14, 3].

In this context of using DNN for decision making, we considered a subset of the informational parameters: Probability of Collision (Pc), Kullback-Leibler (KLD), Mahalanobis Distance (MHD), Bhattacharyya Distance (BD), Miss-Distance (MD), and the Orbit Angle (OA). We grouped the informational parameter assignments of 4 groups, based on their performances in the FIS model:

Group 1 = KLD, MD, BD, Pc, MHD
Group 2 = KL, MD, MHD, Pc
Group 3 = Pc, MHD, OA
Group 4 = Pc

Performance analysis across the four parameter groups revealed that:

Group 3 Pc, MHD, OA achieved the highest performance with the LM algorithm using 10 hidden layers, yielding a regression value of 0.96 and RMSE of 0.0177. This group notably included a geometrical parameter (Orbit Angle) and demonstrated superior error distribution characteristics compared to other configurations. This outcome indicated
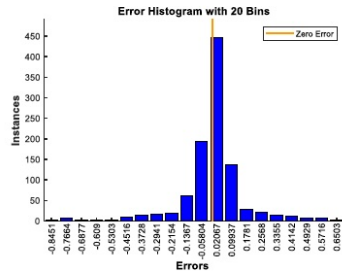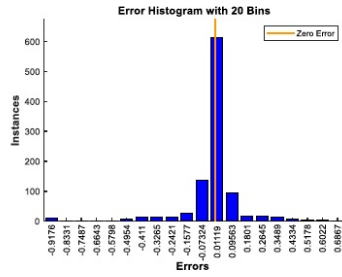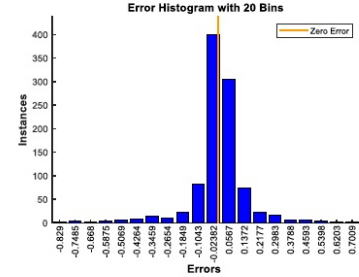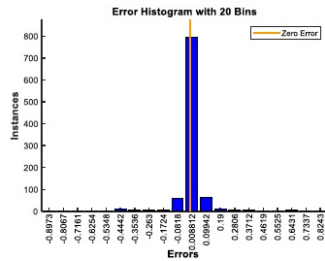
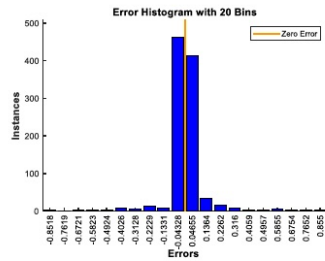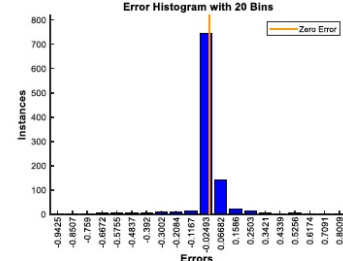Fig. 10: Group 3: Error Histograms for the Scaled-Conjugate Gradien (SCG) (Top Row) and the Levenberg-Marquardt (LM) (Bottom Row) training algortihms for the 10, 20 and 40 hiddden layers investigated[14]

.

that the geometry of the total covariance, or better yet, the progression of the geometry of the total covariance, may provide insight on the prediction accuracies of conjunction risk levels.

Group 1 KLD, MD, BD, Pc, MHD performed optimally with LM training and 40 hidden layers (regression 0.96, RMSE 0.0192), while Group 2 KLD, MD, MHD, Pc showed consistent performance across all layer configurations with the LM algorithm (regression 0.95, RMSE 0.021). Interestingly, Group 4 containing only Pc performed suboptimally despite serving as the target determinant, achieving regression values of 0.93 and RMSE of 0.0303-0.0313, with evidence of overfitting as indicated by training state analysis [14].

The comparative analysis demonstrated that the LM back-propagation algorithm consistently outperformed SCG across all parameter groups, see Figure 10, with Group 3's optimal configuration achieving an overall regression value of 0.958 (training: 0.961, validation: 0.965, testing: 0.934) see Figure 11.
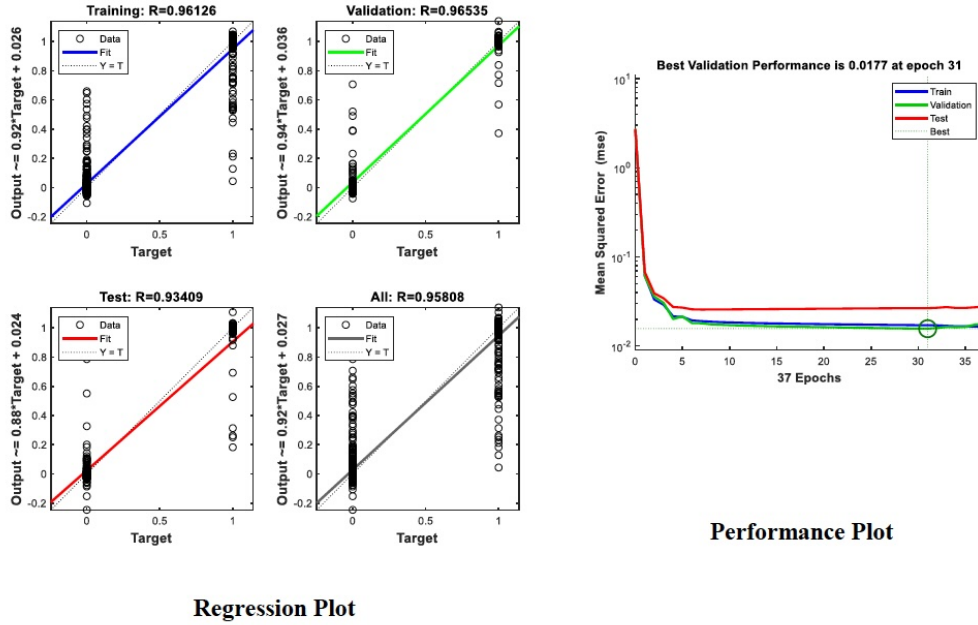

Regression Plot

Performance Plot

Fig. 11: DNN Model performance metrics for Group 3: LM 10 Hidden Layers[14]
.

The study revealed that increasing hidden layers did not necessarily improve the overall performance, and that multi-parameter approaches significantly outperformed single-parameter (Pc-only) models, which mirrors the approach operators take during conjunction assessment and decision support. Focusing the analysis on the single-parameter Pc resulted in overfitting, which signifies the lack of value of using a DNN for single parameter pattern detection analyses. The research concluded that DNNs presented a more promising path than Fuzzy Inference Systems for conjunction assessment information parameter analyses, though further optimization of physics-derived parameter sets and an approach to incorporate future sensor observations for an additional geometrical parameter consideration, would be required, prior to performing a validation run of the model alongside ongoing conjunction assessments in operations.

4.2.4 RNNs: Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) neural network approach was implemented investigate the ability to predict the miss distance with some certainty at TCA using time-series CDM data in advance a finite duration of time. The parameter identification process utilized 18 key parameters including miss distance, relative speed, Energy Dissipation Rate (EDR), Solar Radiation Pressure Coefficient, Mahalanobis Distance, and other various orbital determination metrics [16]. Through unsupervised learning techniques including hierarchical clustering and competitive neural networks, strong correlations were noted between miss distance, probability of collision (Pc), and parameters such as EDR, solar

radiation pressure coefficients, and geomagnetic indices. These clustering methods revealed that perturbation-related parameters (Rcoef, EDR, RCS, AP) were consistently grouped into classes with collision risk metrics across different weight initializations, suggesting their potential importance in miss-distance prediction efficacy.
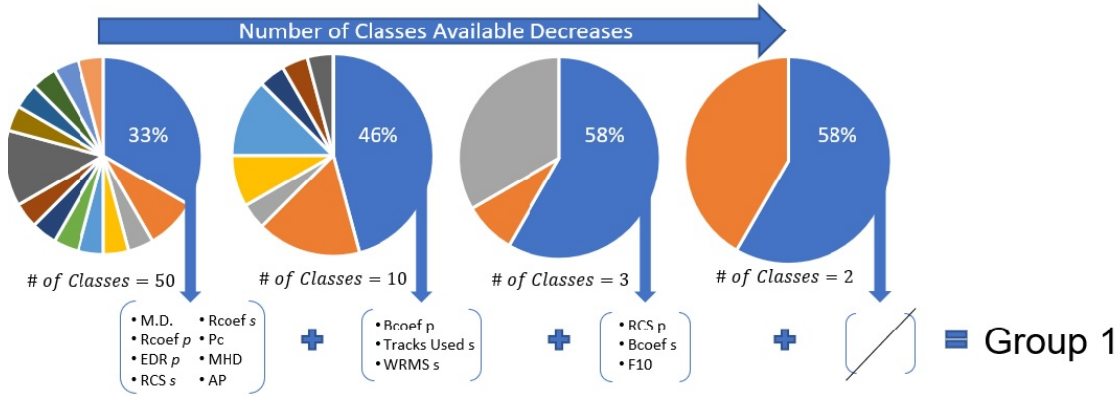

Fig. 12: Competitive Network grouping for close approach events showing strong correlations for the parameters conisdered[16]

.

A shallow binary neural network classifier was implemented to identify events with high-miss distance prediction uncertainty and elevated collision risk. The binary classifier achieved 85.8% overall accuracy for miss distance standard deviation (mdσ) prediction with 85% sensitivity and 86.5% specificity, while attaining 94.4% accuracy with 100% sensitivity for high-risk Pc classification ($Pc > 10^{-8}$). The classifier identified seven critical parameters affecting miss distance predictions: Length of Measurement Update (LUPI), Tracks Used, EDR for both primary and secondary objects, TCA, and Mahalanobis Distance (MHD). The model revealed that below-average MHD values and fewer tracking observations correlated with poor prediction reliability, indicating covariance realism issues in orbital determination.

The model was configured with a batch size of one to enable learning individual event characteristics, and we minimized overfitting through validation of loss monitoring across 100-200 epochs with a learning rate of 0.001 using the Adam optimizer [8]. To improve model robustness, an additive white Gaussian noise layer was applied with μ=0 and σ =0.25 to inputs before processing. Four separate LSTM model groups were constructed with different configurations (as shown in Figure 13) to address the varying nature of conjunction events, each designed to handle the temporal aspects of CDM evolution while accommodating the irregular data availability patterns characteristic of operational conjunction assessment, as shown in Figure 13.

The LSTM architecture incorporated masking layers to handle variable-length CDM sequences, two LSTM layers with ReLU activation, and a time-distributed dense output layer for miss distance prediction. Events with miss distances of less than 5km were used in the training. The optimal Group 2 parameter configuration achieved an average prediction error of 211 meters with a maximum error of 2,480 meters, averaging predictions at 1.8 days before TCA, as shown in Figure 13.

Group	Miss Distance	Relative Speed	LUPI (pri)	Tracks Used (pri)	RCS (pri)	WRMS (pri)	BCoef (pri)	Rcoef (pri)	EDR (pri)	LUPI (sec)	Tracks Used (sec)	RCS (sec)	Bcoef (sec)	Rcoef (sec)	EDR (sec)	Pc	HBR	Time	Angle	MHD	F10	AP	DST
1	x						x	x	x		x	x	x	x	x	x	x			x	x	x	
2	x	x	x	x	x		x	x	x	x	x	x	x	x		x					x	x	x
3	x	x	x	x	x			x	x			x		x						x	x	x	
4	x		x			x				x	x			x		x	x		x		x		

Fig. 13: Groups considered in the LSTM model to determine which parameters minimized miss-distance prediction errors for conjunction events[16]

.

The model demonstrated a strong linear relationship (R^2 correlation) between intrinsic miss distance variance (mdσ) and prediction error, with error approximately doubling for each meter of mdσ increase. Despite data limitations due to the scarcity of close approach events (only 27 high-risk events from 782 total events), the LSTM successfully learned to predict miss distance within a 0.2km accuracy for low-variance events, with performance degrading for events exhibiting high intrinsic miss distance variability across the observation timeline [16], highlighting the stochasticity for conjunction miss-distance predictions. Figure 14 shows the performance of 50 randomly selected CDMs with varying prediction timelines (average of 1.8 days to TCA) utilizing the LSTM model.
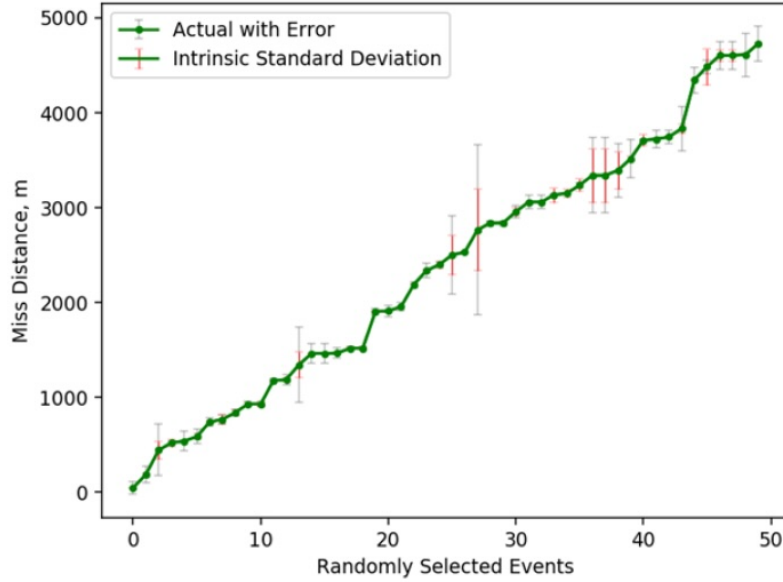

Fig. 14: Group 2 Performance: Prediction of Miss-Distance with respective Standard Deviation (mdσ)[16]
.

5. FINDINGS AND CHALLENGES

The comprehensive investigation of these machine learning approaches revealed several fundamental challenges that limited their operational applicability for conjunction assessment primarily due to stochasticity of the data, the complex and stochastic nature of orbital mechanics, and the difficulty in achieving the high reliability and explainability standards required for a high-value, high-risk space operations decision-making paradigm. Additionally, the "explainability" paradox inherent in many machine learning approaches posed significant concerns for operational implementation, as conjunction assessment decisions require clear justification and understanding of the underlying reasons and factors for each outcome. This research highlighted that while machine learning techniques showed theoretical promise for enhancing conjunction assessment capabilities, significant additional work would be needed to overcome stochastic data limitations, improve model interpretability, and demonstrate reliable performance via validation across the full spectrum of operational conjunction scenarios before these methods could be considered for operational deployment.

5.1 Comparative Analysis of Different AI/ML Approaches

When comparing the various machine learning approaches investigated, clear distinctions emerged in their applicability and performance for conjunction assessment. The unsupervised learning techniques (hierarchical clustering and competitive networks) proved valuable for parameter relationship discovery but inadequate for operational risk classification. These methods successfully identified correlations between parameters such as miss distance and probability of collision, but failed to translate these correlations into actionable risk assessments. Limitations included difficulty in interpreting cluster meanings in operational contexts for repeatable interpretations, challenges in determining optimal

cluster numbers, and the lack of direct connection or optimal determination between identified clusters and actionable risk assessment decisions. In contrast, while the Fuzzy Inference System offered a more interpretable framework that could incorporate expert knowledge through defined parameter thresholds, FIS struggled with establishing reliable weight assignments that outperformed traditional Pc calculations as the ground truth. Despite its theoretical appeal, the FIS approach faced limitations in establishing reliable weight assignments and threshold definitions, as the fuzzy logic rules required extensive domain expertise to construct meaningful decision frameworks that could be operationally implemented deterministically each time.

The supervised learning approaches demonstrated a stark difference between shallow and deep neural network architectures. The shallow binary classification networks showed marginal improvements in risk classification when incorporating multiple information parameters beyond just Pc, with Model 4 (incorporating all available parameters) achieving the best performance metrics [16]. However, the improvement appeared extremely marginal compared to using normal Pc thresholds, raising questions about the practical value of the additional computational complexity. The Deep Neural Networks when applied to static CDM analysis and particularly the LSTM models designed for time-series prediction (multiple CDM time-series data sets), theoretically offered the most promising approach for handling the evolving nature of conjunction data but suffered from insufficient "ground truth data" for training and poor generalization to diverse conjunction scenarios [14].

Each successive model iteration and approach led to increasingly constrained datasets to address specific limitations. The most sophisticated models ultimately relied on the smallest, most constrained datasets, highlighting a fundamental tension between model complexity and data availability in this application domain. The Deep Neural Networks, while theoretically more capable of capturing complex relationships, required substantially more data than was available for reliable training, resulting in poor operational performance compared to current traditional statistical methods.

5.2 Identification of Parameters Associated with High-Risk Events

The challenge of determining optimal parameter combinations and the difficulty in establishing clear thresholds for decision-making, was not a clear undertaking as the relationship between these information parameters and actual collision risk remained complex and are not always intuitive for accurate predictions, due to the unique nature of each conjunction event.

The investigation yielded several insights regarding which parameters showed the strongest association with high-risk conjunction events. Across both supervised and unsupervised approaches, the miss distance parameter (obviously, since this is about collision avoidance - avoiding close proximity) consistently emerged as the most significant parameter correlated with collision risk, followed by probability of collision (Pc), approach angle, and relative velocity. The hierarchical clustering and competitive network analyses revealed strong correlations between miss distance and Mahalanobis distance metrics, suggesting that these statistical measures capture similar aspects of conjunction geometry relevant to risk assessment.

Unsurprisingly, certain combinations of parameters showed stronger predictive power than individual metrics alone. The radial, in-track, and cross-track (RIC) components of relative position and velocity, particularly when analyzed for consistency across time steps, demonstrated significant potential for identifying events with poor miss distance prediction [16, 3]. Events exhibiting high variance in these parameters across successive CDMs were more likely to represent either high-risk scenarios or chattering/fluctuating behavior requiring closer analyst attention. Especially, when considering space weather parameters, particularly the F10.7 solar flux index, showed non-trivial correlations with conjunction risk for satellites in low Earth orbit, highlighting the importance of considering space weather factors and an accurate capture of their complex variability in risk assessment.

However, the research also revealed significant limitations in consistent parameter identification for the overall conjunction assessment. Even though certain parameters showed statistical correlation with high-risk events, these correlations often lacked sufficient strength or consistency to enable reliable operational decision-making. The analysis of parameter importance in the neural network models showed that while adding more parameters did incrementally improve performance metrics, the marginal benefits were minimal beyond the core parameters of miss distance, Pc, and approach geometry. This finding suggests that the complexity added by incorporating numerous additional parameters might not justify the increased computational requirements and reduced model interpretability from an operational perspective.

5.3 Fundamental Challenges and Bottlenecks

The study identified a major challenge: that the stochastic and non-linear nature of the conjunction evolution created inconsistent parameter behavior patterns that were difficult for machine learning models to capture reliably. The temporal analysis revealed that sample cases generated from the same baseline often exhibited completely different behavioral patterns, undermining the assumption that historical patterns could reliably predict future conjunction evolution, within acceptable uncertainty bounds that would account for predicted measurement updates and stochastic perturbations. Additionally, the rarity of actual collision events in the historical dataset created a severe class imbalance problem, making it difficult to train supervised learning models that could accurately distinguish between safe and unsafe encounters.

In particular, the "black box" nature of neural network approaches provides little insight into the physical reasoning behind their predictions, which is a critical need for operational decision-making where understanding the basis for risk assessment is essential.

This effort concluded that significant advances in data collection, model interpretability, and validation methodologies would be required before machine learning techniques could be considered viable for operational conjunction assessment applications.

6. FUTURE DIRECTIONS AND CONCLUSIONS

While AI/ML approaches investigated to date have not yet demonstrated sufficient reliability for operational collision avoidance decision making, several promising directions for future research remain. Although beyond CARA's direct purview, AI/ML techniques could potentially be applied to improve conjunction screening processes through development of adaptive screening volumes that optimize computational resources while maintaining safety margins. As noted in [16], investigating Recurring Neural Networks (RNNs) and LSTM models to ingest time-series based information could eventually provide more reliable predictions at TCA, especially if approaches can be developed that incorporate uncertainty quantification in the predictions. Future studies should carefully consider that predictions of future measurement updates are nearly impossible to model deterministically since they are not entirely physics-based, and that routine maneuvers as part of normal operations are difficult or impossible to model or incorporate within a model. These critical limitations must be addressed by any AI/ML approach for it to be operationally viable.

Recommendations for continued improvements on conjunction assessment processes may focus on addressing fundamental data limitations and physics-based constraints on data quality. A key recommendation involves improving covariance realism, as unrealistic covariance matrices fundamentally undermine the reliability of any risk assessment approach in addition to any machine learning model. Additionally, the development of weighted loss functions that address the class imbalance problem by giving higher importance to the rare high-risk events could improve model training outcomes if implemented carefully. Perhaps most importantly, future work should focus on explainable AI techniques that provide transparency into how models arrive at risk assessments, making them more trustworthy for operational decision making where accountability is essential.

While fully automated AI/ML decision systems are not currently viable for operational risk assessment, hybrid approaches combining traditional physics-based methods with targeted machine learning components may offer incremental improvements. As large constellations continue to launch, the need for more sophisticated automation becomes increasingly urgent, but such automation must maintain the high reliability standards required for space operations. Such as identifying which conjunction events deserve analyst attention, rather than performing risk assessments and conclusive early decision-support for mitigative maneuver decisions. This tiered approach could help address the scaling challenges as the space object catalog grows dramatically in the coming years.

By carefully constraining the problem domain and setting realistic expectations, future AI/ML approaches may yet find valuable niches within the broader conjunction assessment workflow, even if they cannot yet replace traditional physics-based methods entirely.

7. REFERENCES

[1] S. Alfano. Satellite conjunction monte carlo analysis. In *AAS Space Flight Mechanics Meeting, Pittsburgh PA*, 2009.

[2] K. T. Alfriend and M. R. Akella. Probability of collision between space objects,". *Journal of Guidance Control and Dynamics,*, 23(5), 2000.

[3] A Burton, M Zielinski, C Frueh, A Mashiku, , and Memarsadeghi. Assessing measures to reliably predict collisions in the presence of uncertainty. In *28th AAS/AIAA Astrodynamics Specialist Conference, Snowbird, UT*, 2018.

[4] Soumaya Elkantassi, Valérie Chavez-Demoulin, Anthony C. Davison, Matthew D. Hejduk, and Hunter A. Morris. Statistical inference on the miss distance compared to collision probability for conjunction analysis, 2025.

[5] S Gilpin, B Qian, and I Davidson. Efficient hierarchical clustering of large high dimensional datasets. In *Proceedings of the 22nd ACM international conference on Information Knowledge Management*, pages 1371–1380, 2013.

[6] M.T Hagan and M Menhaj. Training feed-forward networks with the marquardt algorithm. *IEEE Transactions on Neural Networds*, 5(6):989–993, 1994.

[7] D.T Hall, S.J Casali, L.S Johnson, B.B. Skrehart, and L.G. Baars. High fidelity collision probabilities estimated using brute force monte carlo simulations,". In *AIAA/AAS Astrodynamics Specialist Conference, Snowbird, UT*, 2018.

[8] S Hochreiter and J Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[9] The Mathworkds Inc. Introducing deep learning with matlab. In *Deep Learning eBook*, 2017.

[10] The Mathworkds Inc. Build fuzzy systems using fuzzy logic designer. In *https://www.mathworks.com/help/fuzzy/building-systems-with-fuzzy-logic-toolbox-software.html,*, 2019.

[11] Darpa: Explainable Artificial Intelligence. Explainable artificial intelligence. In *https://www.darpa.mil/research/programs/explainable-artificial-intelligence*, 2024.

[12] D Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.

[13] A Mashiku. A compendium for artificial intelligence and machine learning for satellite collison avoidance at nasa conjunction assessment risk analysis (cara). In *Cognitive Space Operations (COGSPO)*, 2024.

[14] A Mashiku, C Frueh, N Memarsadeghi, E Gizzi, M Zielinski, and A Burton. Predicting satellite close approaches using statistical parameters in the context of artificial intelligence. In *29th AAS/AIAA Astrodynamics Specialist Conference, Portland, ME*, 2019.

[15] Office of the Chief Engineer. Nasa spacecraft conjunction assessment and collision avoidance best practices handbook. 1, 2023.

[16] B.I Robertson and A Mashiku. Early information parameter-set analysis for satellite close approaches using machine learning. *Icarus*, 62(2):244–272, 1985.

[17] D.F Specht. A general regression neural network. *IEEE Transactions on Neural Networks*, 2(6):568–576, 2000.

[18] M Steinbach, G Karypis, and V Kumar. A comparison of document clustering techniques. In *University of Minnesota, Technical Report 00-034,*, number 00-034, 2000.

[19] A Tripathi. What is the main difference between rnn and lstm. In *Learn Data Science, Machine Learning and Artifical Intelligence*, 2021.

[20] F Vasquez. *Deep Learning made easy with Deep Cognition*. Becoming Human: Aritficial Intelligence Magazine, Dec 21, 2017.

[21] H Yu and B.M Wilamowski. *Levenberg-Marquardt Training*. Auburn University, 2010.