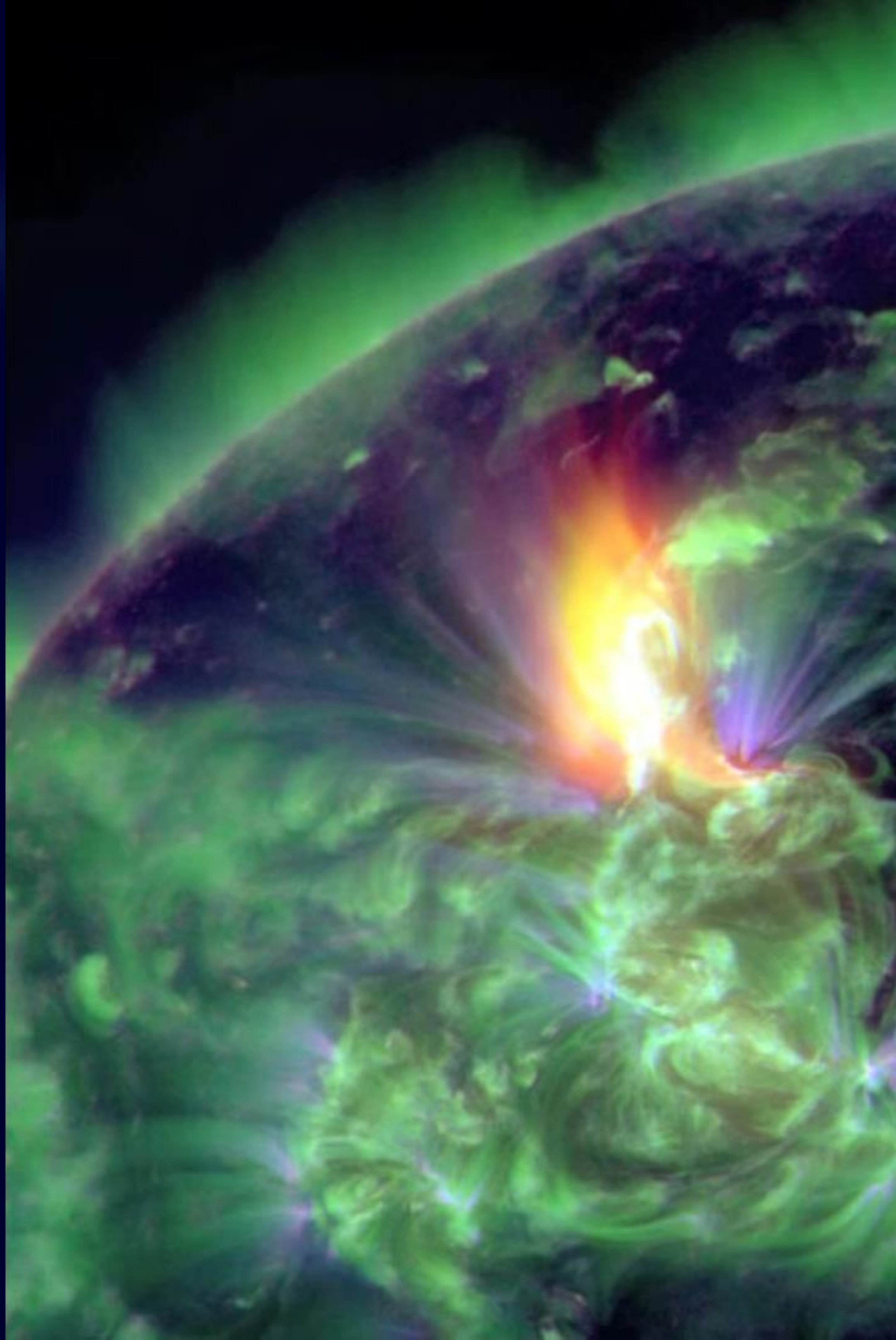
Machine Learning Model to Estimate the Damage on the Economy due to Solar Storm over a Specific Region

by
Mili Chowdhury
Pratishruti Sahoo
Subhadeep Dey

Introduction

Solar storms can have devastating effects on modern infrastructure and pose serious risks to our interconnected world.

This presentation will explore the use of machine learning models to accurately assess and mitigate the potential damage from solar activity in a specific region on the Earth and finally their effects in the economy.



Uniqueness of the Model

1 Estimation of the Damage

This model forecasts the damage by taking into account how susceptible the infrastructure is to potential solar storms.

2 Geospatial Adaptability

In addition, it forecasts the monetary loss for a given area based on an assessment of the potential solar storm's damage.

3 Proactive Approach

These two factors allow the model to provide information about the preventative steps that should be taken in case of a solar storm over that area.

Effects of Solar Storm on the Earth

1

Electrical Grid Disruptions

Solar storms can induce powerful currents in power lines, causing widespread blackouts and damage to electrical infrastructure.

2

Satellite & Communication Disruptions

Solar radiation can interfere with satellite operations, GPS signals, and communication networks, hampering modern connectivity.

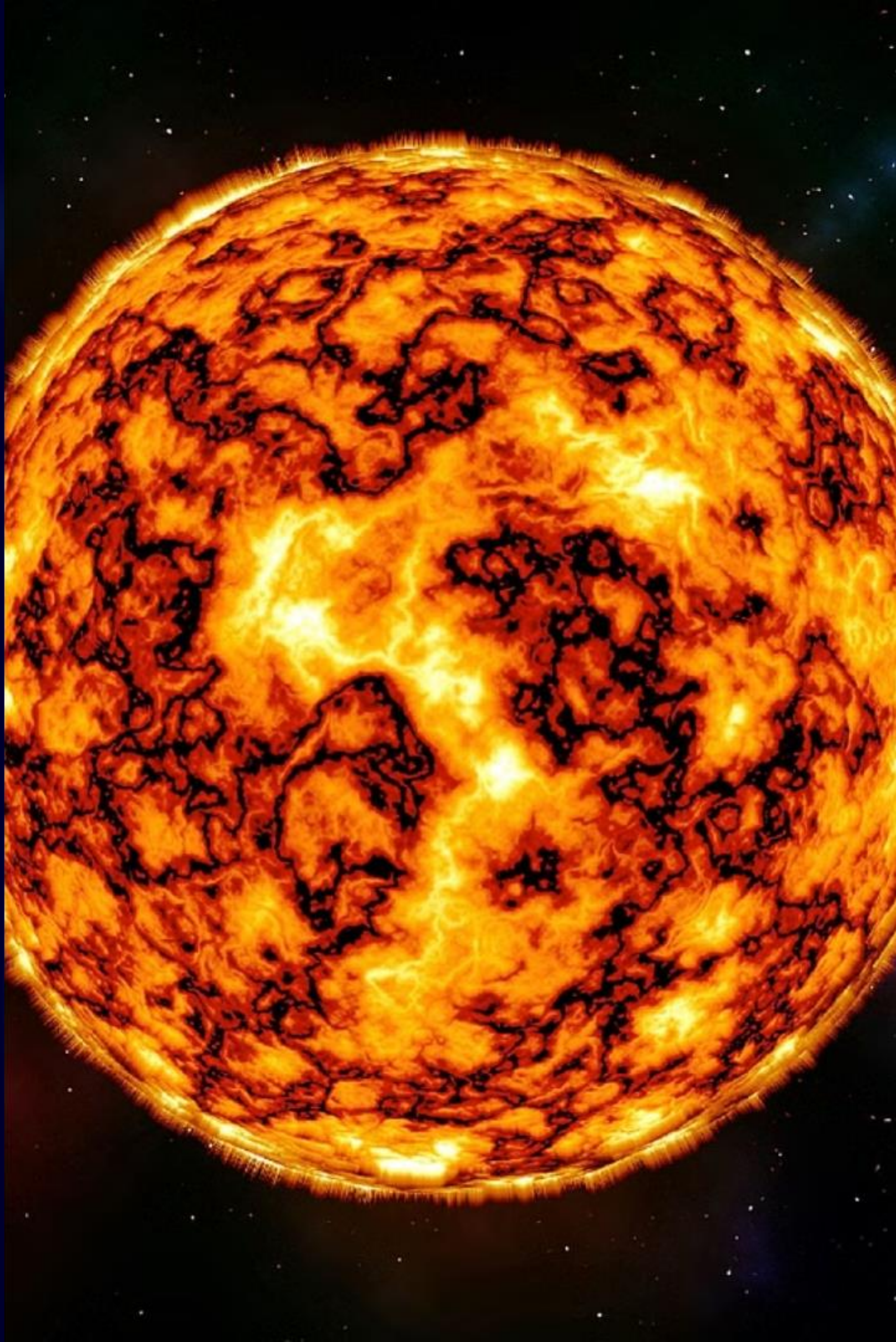
3

Airline & Navigation Disruptions

Increased radiation during solar storms can force airlines to reroute flights, leading to delays and cancellations due to interruption on RADARs. Navigation systems can also be affected.

Previous Rampages of the Sun

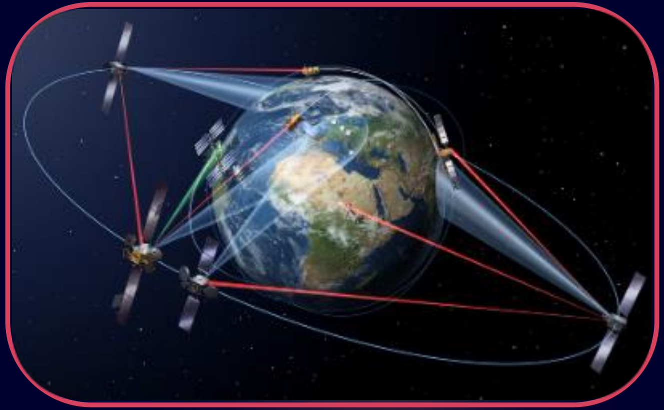
1. The Carrington Event (September 1859)
2. The Bastille Day Solar Storm (July 2000)
3. The Halloween Solar Storms (Oct 2003)



Practical Applications of This Model

Power Grid Resilience

Machine learning models can help to predict and can try mitigate the impact of solar storms on power grid infrastructure, enabling proactive risk management and faster restoration of services.



Satellite Communications

These models can also assess the vulnerability of satellite communications networks to solar storm disruptions, informing system design and backup planning for mission-critical applications.

GPS and Navigation

By forecasting the effects of solar storms on GPS and navigation systems, these ML models can support reliable transportation, aviation, and location-based services during extreme space weather events.

Challenges in Building the Model

Paucity of Detailed Historical Records

Numerous incidents that occurred throughout history are not documented.

Non-Centralized Data

The different helioscopic data are recorded by different observatories (both on the Earth and satellites).

Geospatial & Demographic Data

Details about the population and asset locations in a particular area that the solar storm is expected to affect.

Integrating Data regarding Disaster-prone Infrastructures

Detailed information on the power grid infrastructure, including the locations of substations and transmission lines, along with the data vis-a-vis the damage on the satellites, telecommunication networks etc.

Data Collection and Preprocessing

1. Gather historical data on solar storms and their impact on power grids, communication systems, and satellites in the target region.
2. Collect detailed geospatial data such as population density, and asset locations, in order to model the vulnerability of the region.
3. Preprocess the data to handle missing values, outliers, and inconsistencies, ensuring data quality and integrity for the machine learning model.

Rationale behind using ML over Statistical Analysis

1 Nonlinearity Modeling

The relationship between geomagnetic disturbance and solar activity is nonlinear. Additionally, it is dynamic and changes rapidly over time. ML is superior because it also requires the integration of data from many sources.

2 Handling Uncertainty

Machine learning models can handle large and diverse datasets and can adopt over time. Models like semi-supervised and reinforcement machine learning can work with sparse data and provide accurate predictions.

3 Scalability and Adaptability

Such a model like ours involve complete high dimensional data and non-linear relationships. Our data also can face multi-collinearity and autocorrelation which a machine learning model can handle better than a statistical model.

4 Compilation of Data

Such a model involves complex data from various resources such as Solar absorptions, Geomagnetic measurements, and historical solar storm data. Such variables may not be easily discernible through tradition statistical methods.

Feature Engineering and Selection

1 **Data Exploration**

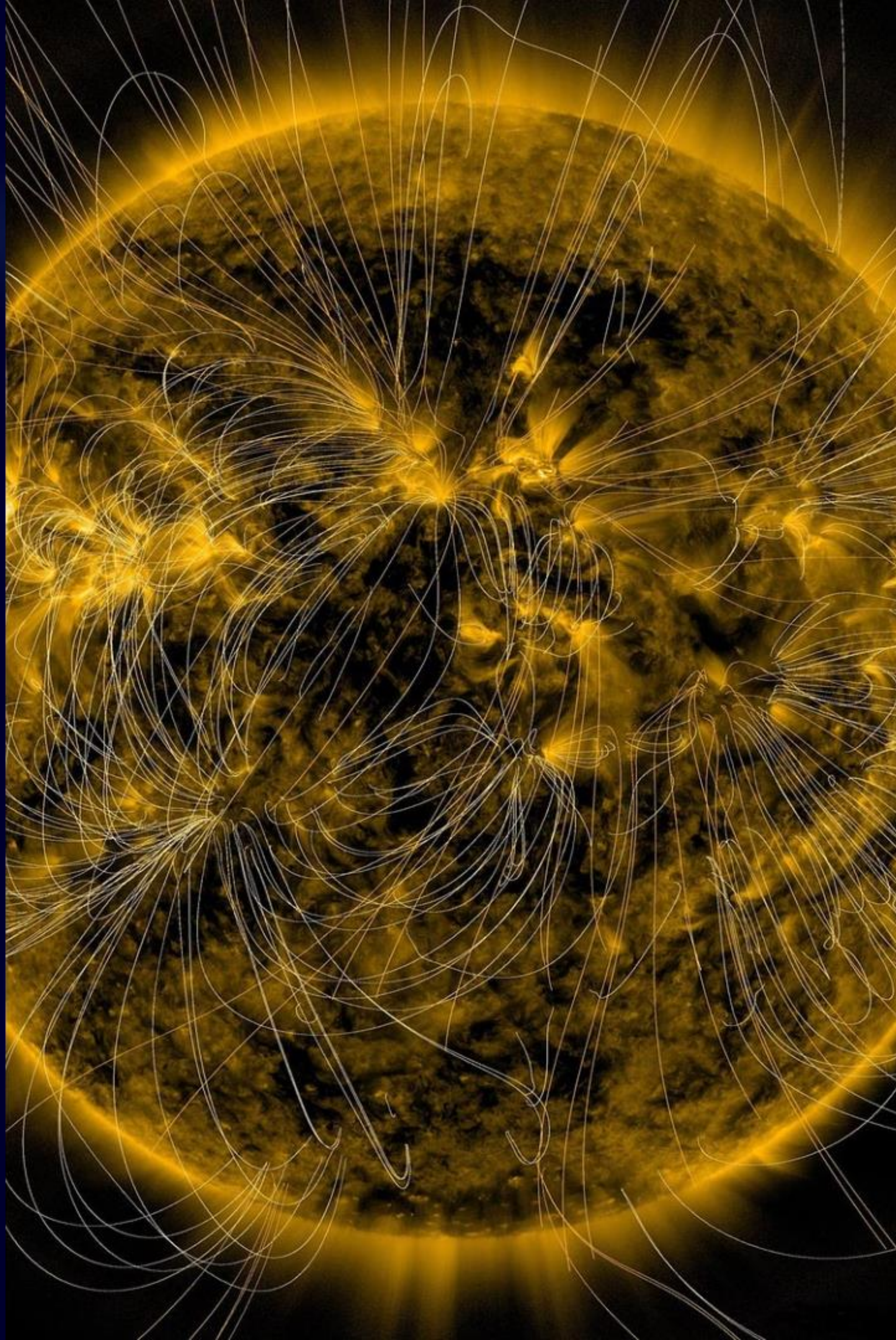
Conduct a thorough analysis of the available data, including statistical summaries, visualizations, and identification of key variables that may influence solar storm damage.

2 **Feature Generation**

By integrating or modifying the raw data, one can produce new characteristics like classifying the damages into distinct groups according to the damages within different sectors.

3 **Feature Selection**

Use advanced techniques like correlation analysis, recursive feature elimination to identify the most informative and predictive features for the damage estimation model.



Damage Estimation Methodology

1

Data Modeling

Develop predictive models to estimate damage based on historical data and simulations.

2

Vulnerability Assessment

Analyze infrastructure susceptibility and exposure to solar storm impacts.

3

Risk Quantification

Quantify potential economic losses and disruptions.

Damage Estimation Methodology

The damage estimation methodology involves a multifaceted approach.

First, developing data-driven predictive models to estimate the extent of damage based on **historical observations** with the help of regression tree algorithm.

Second, conducting **a detailed vulnerability assessment** to understand the susceptibility and exposure of critical infrastructure within the region, and quantifying the potential economic losses, service disruptions that could result from a severe solar storm.

Finally, **in order to adjust the inflation rate**, a base year will be taken and as per the present economic conditions the predicted value of economic loss will be determined.

Dummy Dataset

Raw Data

Year	Countries/ Regions	Possible affected area (lakh sq km)	Intensity of the solar storm (W/m^2)	Severity	Possible affected population (lakhs)	Possible Effecting Power Grid (MW)	Possible Damage in Power Grid (MW)	Possible Effecting Surface Telecome Network (MHz)	Possible Damage in Telecom/surface telecom network (MHz)	Possible Effects in Satellites	Possible Damage in Satelites	Total monetary loss ($)
1859	Australia	14	10	Moderate	210	-	-	-	R1	-	-	200
1872	UK, Denmark	11	100	Minor	165	70	89	R2	R2	-	-	100
1921	USA	22	1000	Strong	330	130	180	R3	R3	-	-	300
1938	North America	109	100	Minor	747	110	96	R2	R2	-	-	400
1989	USA, Canada	102	10000	Severe	255	155	190	R4	R4	G1	G2	900
2003	Sweden	23	100000	Extreme	345	217	200	R5	R5	G3	G3	1000

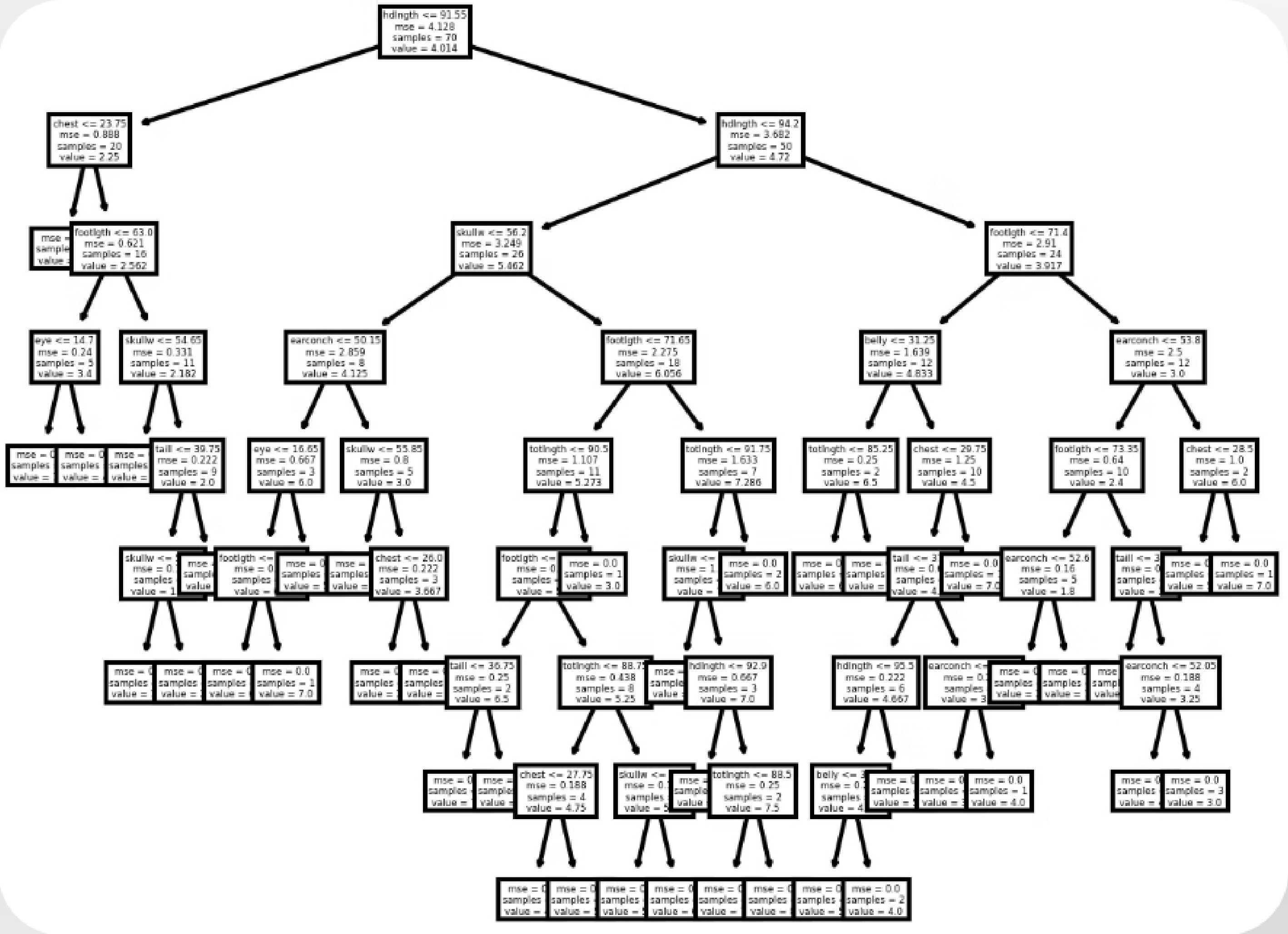
Dataset with Selected Features

Year	Countries/ Regions	Possible affected area (lakh sq km)	Intensity of the solar storm (W/m^2)	Severity	Possible affected population (lakhs)	Possible Effecting Power Grid (MW)	Possible Damage in Power Grid (MW)	Possible Effecting Surface Telecome Network (MHz)	Possible Damage in Telecom/surface telecom network (MHz)	Possible Effects in Satellites	Possible Damage in Satelites	Total monetary loss ($)
1859	Australia	14	10	Moderate	210	-	-	-	R1	-	-	200
1872	UK, Denmark	11	100	Minor	165	70	89	R2	R2	-	-	100
1921	USA	22	1000	Strong	330	130	180	R3	R3	-	-	300
1938	North America	109	100	Minor	747	110	96	R2	R2	-	-	400
1989	USA, Canada	102	10000	Severe	255	155	190	R4	R4	G1	G2	900
2003	Sweden	23	100000	Extreme	345	217	200	R5	R5	G3	G3	1000

Model Architecture and Training

The machine learning model for solar storm damage estimation utilizes the **regression tree** algorithm. This model can capture the relationships between various input features and the predicted damage levels.

The model is trained on a dataset of historical solar storm events and observed damage patterns.



First Three Regression Trees

Possible affected area (lakh sq km)	Intensity of the solar storm (W/m^2)	Possible affected population (lakhs)	Possible Effecting Power Grid (MW)	Possible Damage in Power Grid (MW) [TARGET]
14	10	210	-	-
11	100	165	70	89
22	1000	330	130	180
109	100	747	110	96
102	10000	255	155	190
23	100000	345	217	200

Regression Tree 1

For estimating the damage on the power grid resilience

Possible affected area (lakh sq km)	Intensity of the solar storm (W/m^2)	Possible affected population (lakhs)	Possible Effecting Surface Telecome Network (MHz)	Possible Damage in Telecom/surface telecom network (MHz) [TARGET]
14	10	210	-	R1
11	100	165	R2	R2
22	1000	330	R3	R3
109	100	747	R2	R2
102	10000	255	R4	R4
23	100000	345	R5	R5

Regression Tree 2

For estimating the damage on the telecom network on the Earth's surface

Possible affected area (lakh sq km)	Intensity of the solar storm (W/m^2)	Possible affected population (lakhs)	Possible Effects in Satellites	Possible Damage in Satelites [TARGET]
14	10	210	-	-
11	100	165	-	-
22	1000	330	-	-
109	100	747	-	-
102	10000	255	G1	G2
23	100000	345	G3	G3

Regression Tree 3

For estimating the damage on the space satellite network

Final Regression Tree

	A	B	C	D
1	Possible Damage in Power Grid (MW)	Possible Damage in Telecom/surface telecom network (MHz)	Possible Damage in Satelites	Total monetary loss ($) [TARGET]
2	-	R1	-	200
3	89	R2	-	100
4	180	R3	-	300
5	96	R2	-	400
6	190	R4	G2	900
7	200	R5	G3	1000
8				

Final Regression Tree
For estimating the total monetary loss

Model Evaluation and Validation

Evaluation Metrics

Carefully select appropriate evaluation metrics to assess the model's performance, such as **accuracy** and **root mean squared error (RMSE)** for regression tasks. *It should also be noted that using RMSLE will be preferable if the predicted values' range is approaching a large value.* (N.B.: While calculating the accuracy, the predicted value will be considered as a correct value if it is within the range of 5% on both sides)

Cross-Validation

Employ robust cross-validation techniques to ensure the model's generalization capability and avoid overfitting, such as k-fold cross-validation.

Holdout Testing

Set aside a dedicated testing dataset to evaluate the model's performance on unseen data and gain insights into its real-world applicability.

Sensitivity Analysis

Conduct sensitivity analysis to understand the model's behavior and identify the most influential features, helping to refine the model and feature engineering process.

Sensitivity Analysis and Uncertainty Quantification

1

Sensitivity Analysis

Assess the impact of input parameter variations on model outputs. Identify the most influential factors driving solar storm damage estimates.

2

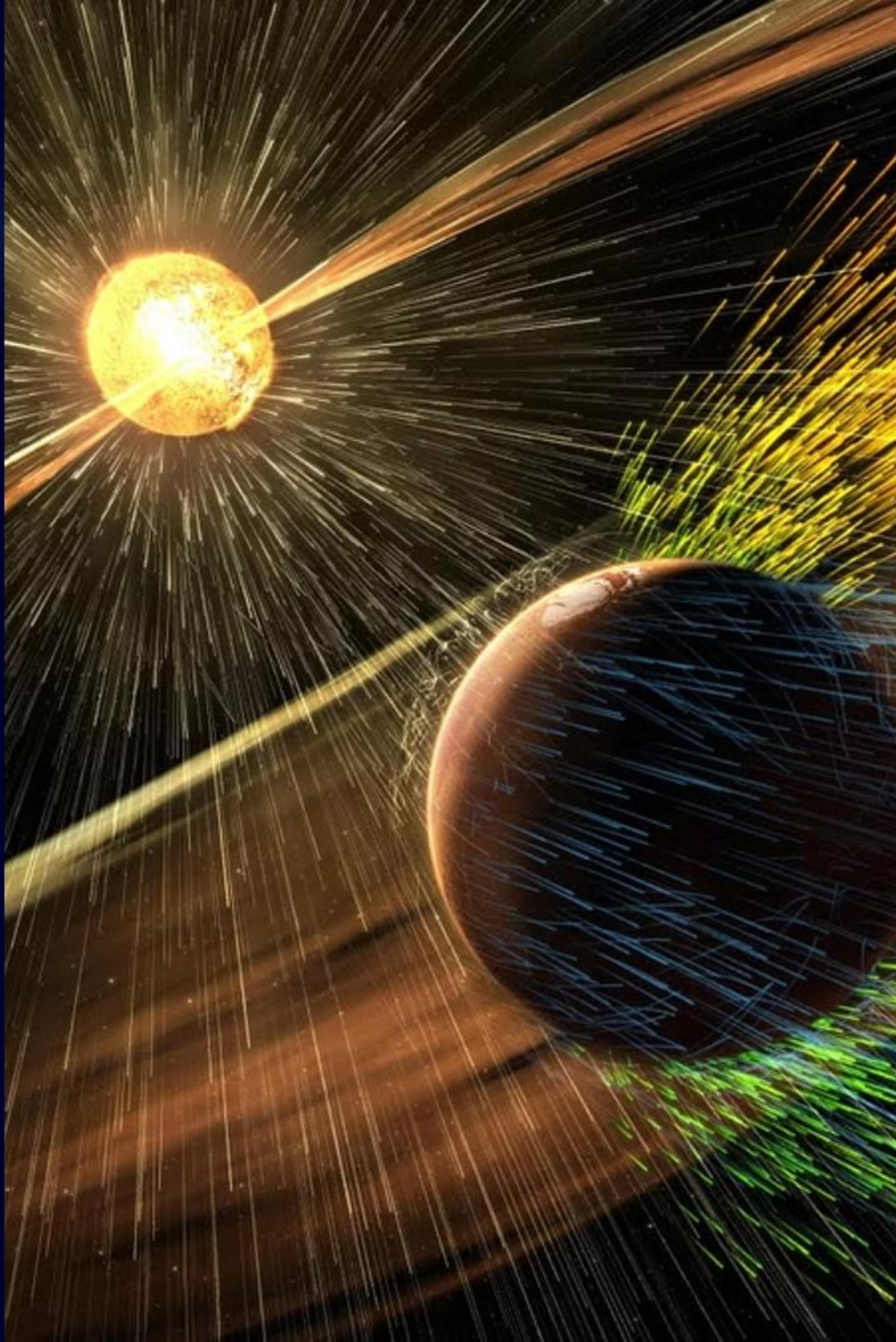
Uncertainty Quantification

Characterize the uncertainty in the input data, model parameters, and assumptions. Propagate these uncertainties through the model to derive probabilistic damage estimates.

3

Scenario Analysis

Simulate multiple plausible solar storm scenarios to understand the range of possible outcomes and their likelihood. Inform risk-based decision making.



Conclusions

Key Takeaways

The machine learning model developed can accurately estimate the economic damage caused by solar storms, enabling proactive risk mitigation strategies.

Global Implications

Solar storms pose a significant threat to modern infrastructure, with potential for widespread disruption. This model can help inform global preparedness efforts.

Outlook

Continued research and refinement of the model can enhance its accuracy and applicability, strengthening our resilience to the impacts of solar storms.

Scope of Improvement

1

Closer to Reality

By adding more characteristics, such as, **weather data**, **geomagnetic data** etc., we can get the model closer to reality and better estimation, but in such a way that the model complexity does not compromise the performance.

2

Expanding Geographical Scope

Future research should explore expanding the model to other regions prone to solar storm activity, allowing for **more comprehensive risk assessment and preparedness planning** across the globe.

3

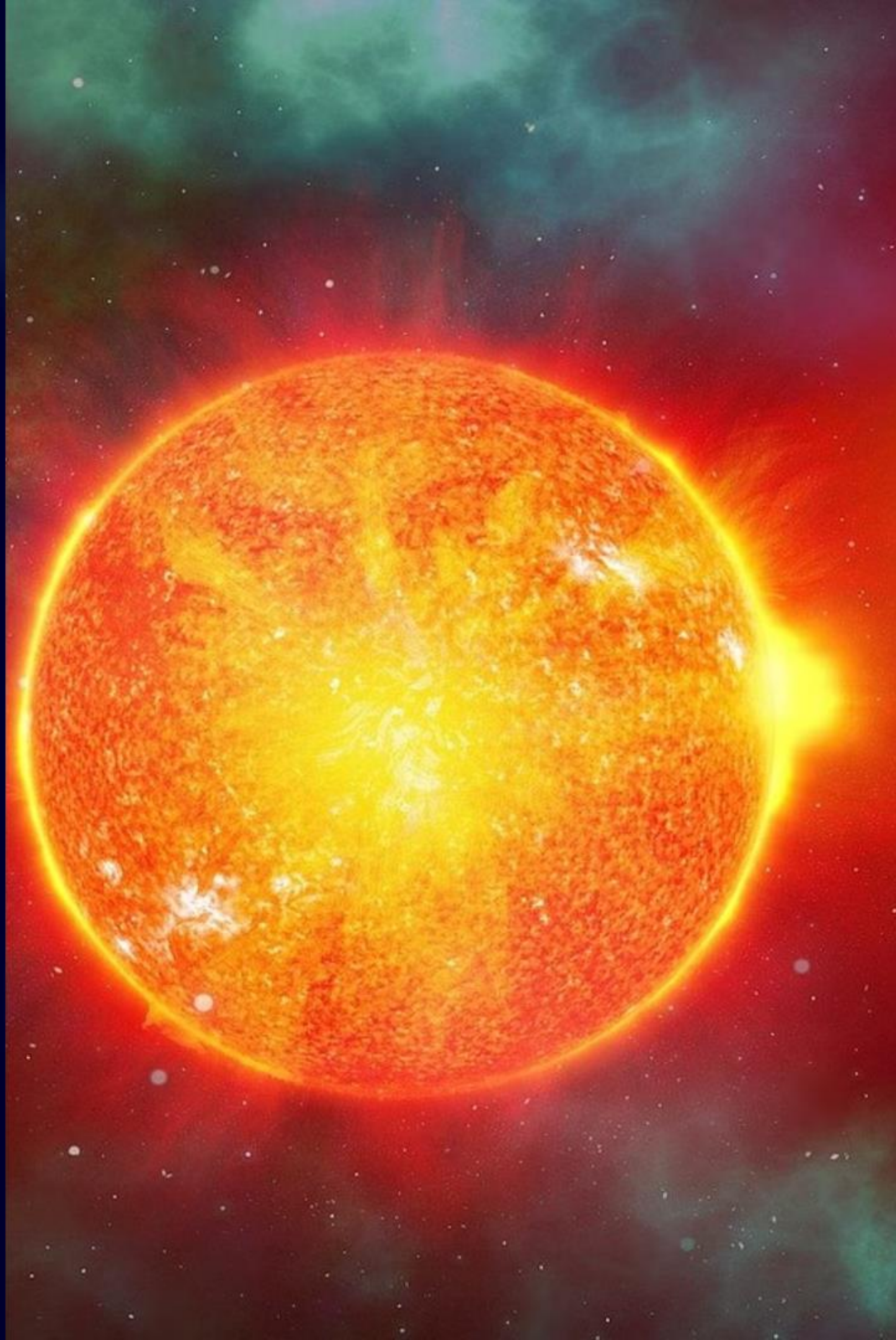
Integrating Real-time Data

Incorporating **live satellite monitoring** and forecasting data could further enhance the model's accuracy and enable near-real-time damage estimation, facilitating more timely emergency response and mitigation strategies.

4

Exploring AI-Powered Simulations

Advancing the model to **leverage cutting-edge AI** and **simulation techniques** could unlock even more sophisticated scenario analysis and predictive capabilities, empowering decision-makers to proactively address the challenges posed by solar storms.



Thank You