New capabilities for ionospheric prediction:

A novel approach with MACHINE LEARNING
How is CEDAR evolving and why do we need data science?

Why is ionospheric scintillation a fantastic use case and what progress have we made?

What trends does this reveal?
How is CEDAR evolving and why do we need data science?
Opportunity:
Evolve traditional approaches
Embrace data science-driven discovery
Enable interdisciplinary work
Scalable architectural approaches, techniques, software and algorithms which alter the paradigm by which data are collected, managed and analyzed.

Dan Crichton, JPL
Someone or something that doesn't fit within traditional academic discipline—a field of study with its own particular words, frameworks, and methods

*Joi Ito, MIT Media Lab, “Antidisciplinary”*
Why is ionospheric scintillation a fantastic use case and what progress have we made?
Global Navigation Satellite System (GNSS) signals for Space Science
Global Navigation Satellite System (GNSS) signals for Space Science

Total electron content (TEC)

Ionosphere

1000 km

100 km
Global Navigation Satellite System (GNSS) signals for Space Science

1000 km

100 km

Ionosphere

Scintillation (signal disruption)

5/16/18

McGranaghan: A new frontier
Global Navigation Satellite System (GNSS) signals for Space Science

- GPS
- GLONASS
- Galileo
- Beidou

Ionosphere

1000 km
100 km
Global Navigation Satellite System (GNSS) signals for Space Science

- GPS
- GLONASS
- Galileo
- Beidou

Support Vector Machine (SVM) | Decision Trees | Random Forests | Neural Networks
---|---|---|---
Easily explainable | | Difficult to explain | |

Create a narrative of new scientific understanding across spectrum of machine learning approaches

Ionosphere

1000 km

100 km
Global Navigation Satellite System (GNSS) signals for Space Science

GPS
GLONASS
Galileo
Beidou

Support Vector Machine (SVM) | Decision Trees | Random Forests | Neural Networks
---|---|---|---
Easily explainable | | | Difficult to explain

Create a narrative of new scientific understanding across spectrum of machine learning approaches

1000 km Ionosphere
100 km
Step 1:
Obtain solar, geomagnetic, and ionospheric data
Step 1:
Obtain solar, geomagnetic, and ionospheric data
Step 1:
Obtain solar, geomagnetic, and ionospheric data
Step 2:
Define the predictive task

GNSS Scintillation
(signal disruption)
Step 3:
Machine learning algorithm for prediction

Support Vector Machine

Electron flux

Interplanetary magnetic field

Cortes and Vapnik (1995)
Step 3:
Machine learning algorithm for prediction

Support Vector Machine

Cortes and Vapnik (1995)
Step 3:
Machine learning algorithm for prediction

Support Vector Machine

Cortes and Vapnik (1995)
Step 3:
Machine learning algorithm for prediction

Support Vector Machine

Cortes and Vapnik (1995)
Step 3:
Machine learning algorithm for prediction

True label

no scintillation

scintillation

Predicted label

no scintillation

scintillation

True negative

False positive

False negative

True positive

McGranaghan et al., (2018)
Step 3: Machine learning algorithm for prediction

McGranaghan et al., (2018)
Step 3:
Machine learning algorithm for prediction

91%
Improved ability to predict when scintillation will occur indicates potential of approach

McGranaghan et al., (2018)
Step 3:
Machine learning algorithm for prediction

67%
High accuracy predicting when scintillation would not occur
Step 4:
Interrogate the model
Evaluation

Step 4:
Interrogate the model
Evaluation

**True Skill Statistic (TSS)**

\[
TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}
\]

**Step 4:**
Interrogate the model

- **True label**
  - **True negative**
  - **False negative**
  - **False positive**
  - **True positive**

- **Predicted label**
  - **no scintillation**
  - **scintillation**
Evaluation

True Skill Statistic (TSS)

\[
TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}
\]

Step 4:
Interrogate the model
Evaluation

True Skill Statistic (TSS)

\[ TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN} \]

-1  Worst

Previous state of the art

SVM benchmark

+1  Perfect

True label

no scintillation

scintillation

True negative

False negative

False positive

True positive

Predicted label

Step 4:
Interrogate the model

Worst

Perfect

SVM benchmark

Previous state of the art
Step 4:
Interrogate the model

Evaluation

Explanation
Step 4:
Interrogate the model

Evaluation

Explanation
Step 4:
Interrogate the mode

Evaluation

Explanation

F-Score (i.e., importance)
Step 4:
Interrogate the mode

Evaluation

Particle precipitation more informative than SW parameters

Explanation

Step 4: Interrogate the mode

Evaluation


Step 4: Interrogate the mode

Evaluation

Explanation
Evaluation

Explanation

SVM predictions $\Phi$

Persistence predictions $\Phi$

True $\Phi$

January 20, 2016

$B_Z$

$B_Y$

$K_p$

$A_E$

$\text{TN}$ $\text{FP}$

$\text{FN}$ $\text{TP}$
January 20, 2016

Evaluation

Explaination

SVM identifies strong scintillation, persistence does not.
Evaluation

January 20, 2016

B<sub>Z</sub>  B<sub>Y</sub>

Kp

AE

SVM predictions σ<sub>φ</sub>

Persisting predictions σ<sub>φ</sub>

True σ<sub>φ</sub>

FP  TN  TP  FN

SVM contains high number of ‘false alarms’
What trends does this reveal?
What trends does this reveal?

Be antidisciplinary
Be antidisciplinary

Be open by default

What *trends* does this reveal?
Be antidisciplinary

Be open by default

Understand the models

What trends does this reveal?
Trends

Be _antidisciplinary_

Be open by default

Understand the models


Curated Sources of Data Science Learning Resources

Ryan McGranaghan running list of resources (Github repository)
  • https://github.com/rmcgranaghan/data_science_tools_and_resources

HelioAnalytics website and list of resources
  • https://sites.google.com/view/heliodata/resources?authuser=0