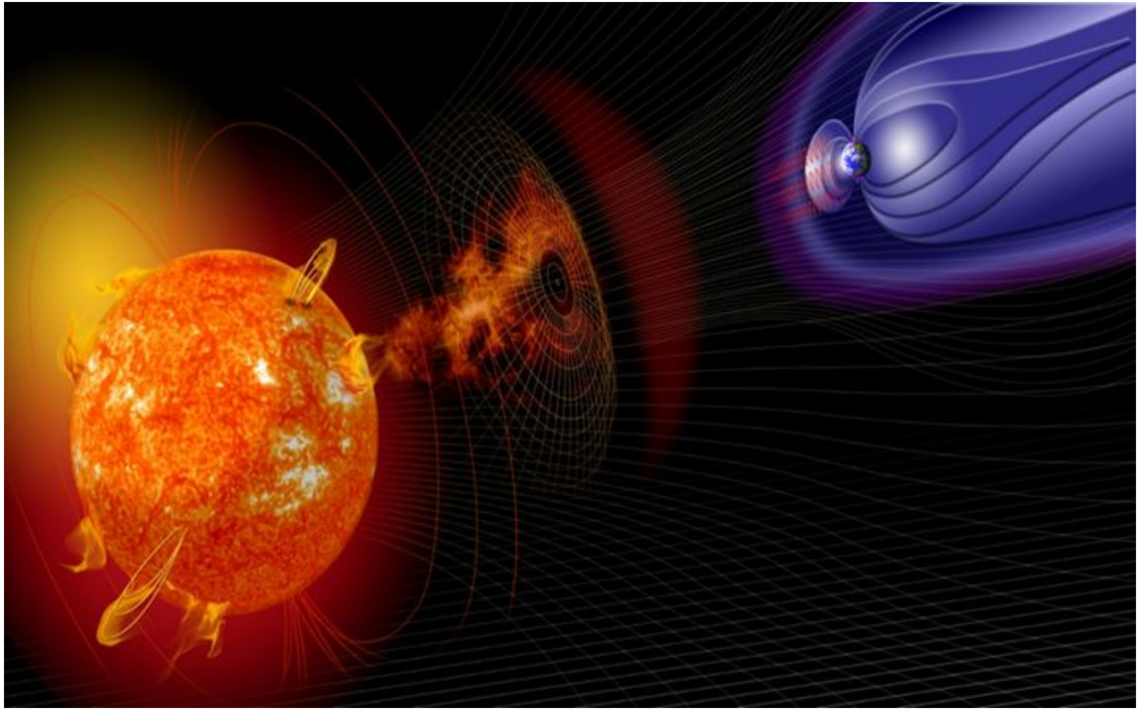
Workshop on Applications of Artificial Intelligence and Machine Learning

Explainable AI in Space Weather Predictions

Hemapriya R (phd1901121008)
Supervisor: Dr.Saurabh Das
Department of Astronomy, Astrophysics and Space Engineering
IIT Indore
Date: 10th Nov, 2022

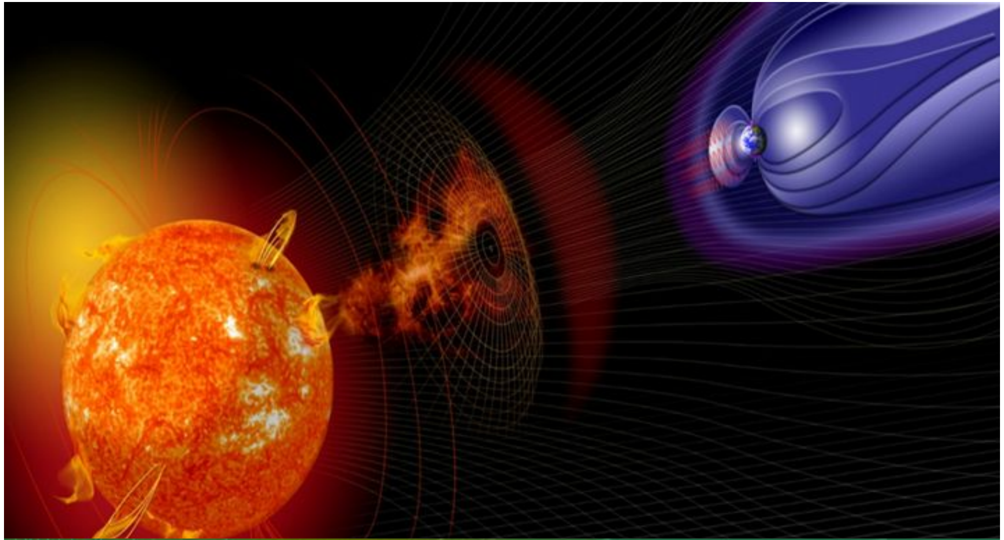
Space Weather - A short Intro

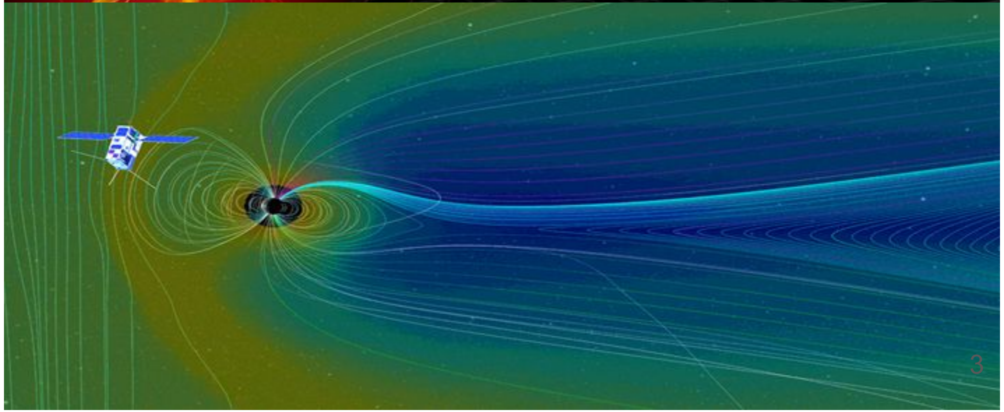
- Akin to terrestrial weather, space weather results from a complex system driven by the Sun and the consequent events much closer to Earth.



Space Weather - A short Intro

- Akin to terrestrial weather, space weather results from a complex system driven by the Sun and the consequent events much closer to Earth.
- The stream of plasma ejected out from the sun travels through interplanetary space.
- Interacts with earth’s magnetic field causing minor to major variations from aurora to Geomagnetic storms.





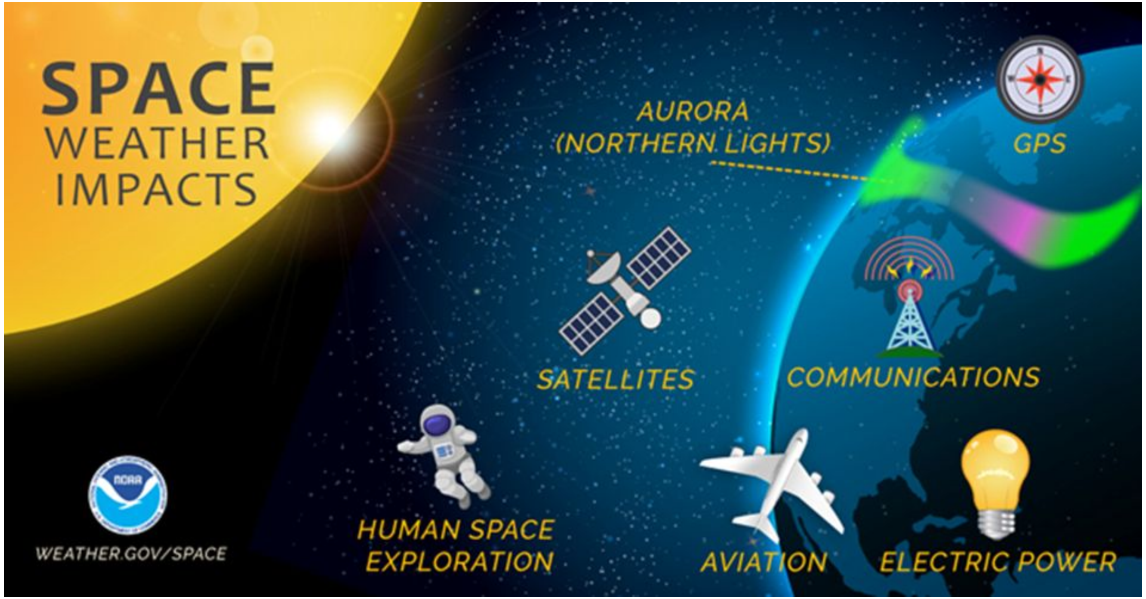
Need for Space weather Predictions

- Evaluate the nature of Geomagnetic storms in advance

- Safe human space exploration

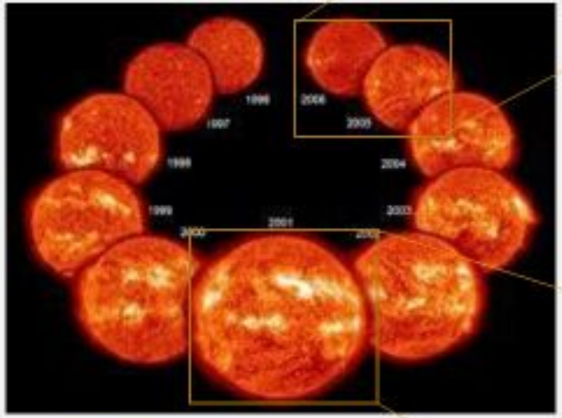
- To avoid major power outages

- To avoid disturbances in HF communication, as ionosphere density varies with storms and radiation



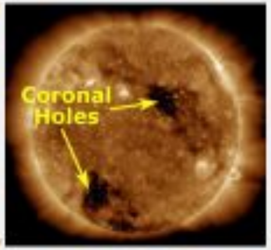
SOURCE OF SPACE WEATHER EVENTS

11 year solar cycle



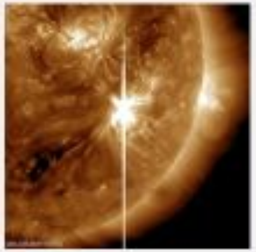
Source: https://www.sansa.org.za

Coronal Holes



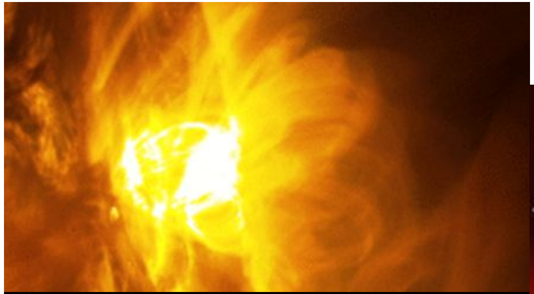
- Prominent in Descending phase
- Speed 300-1000 Km/s
- Reaches earth 2 to 5 days
- Recurrent and stable

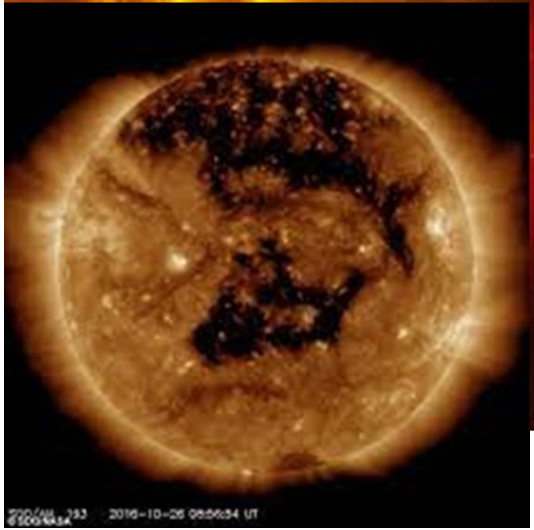
Coronal Mass Ejections(CMEs)

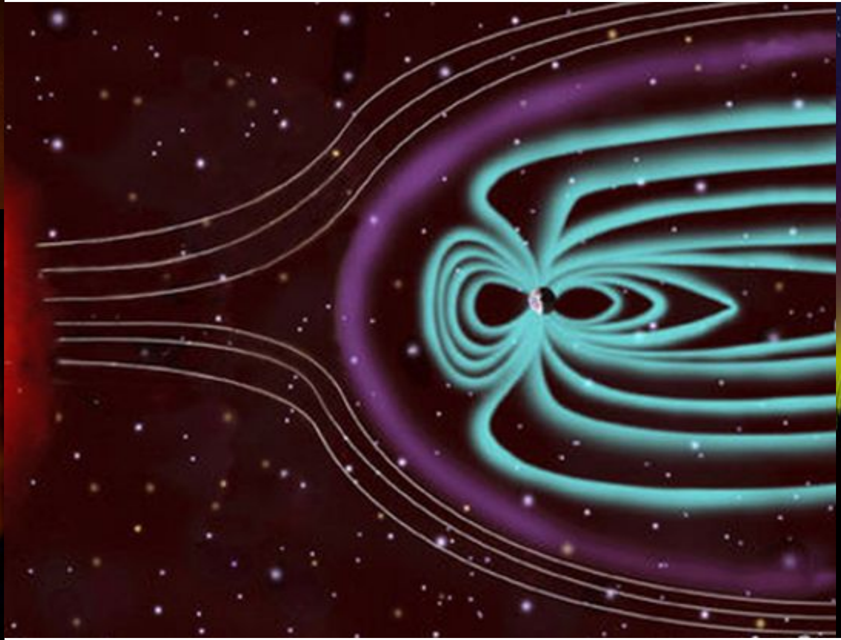


- Prominent in Solar maximum
- Speed 500-1500 Km/s
- Reaches earth within 15 hrs to 3 days
- Sudden and intense

Space weather Prediction : Solar wind speed







Solar wind speed prediction: Data Source

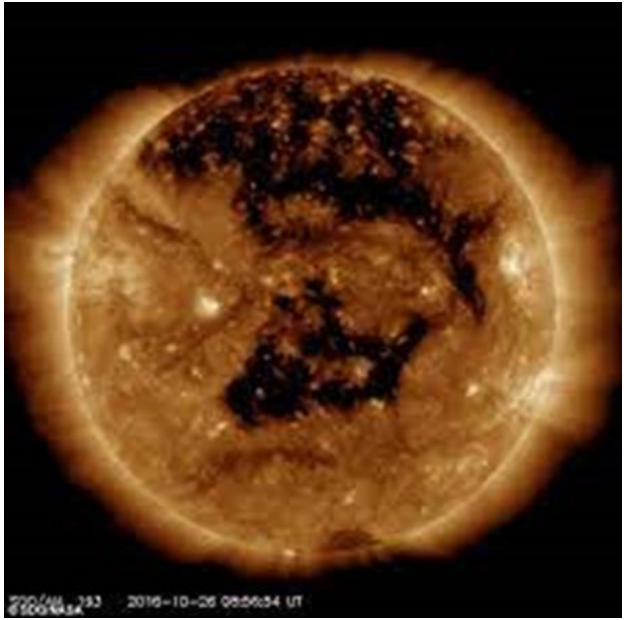
Source:

Input dataset:

- SDO/AIA -193Å(FITS file)
- Spatial Resolution: 512x512
- Temporal Resolution: 2 hr
- Link:Stanford ML curated dataset
- https://purl.stanford.edu/jc488jb7715

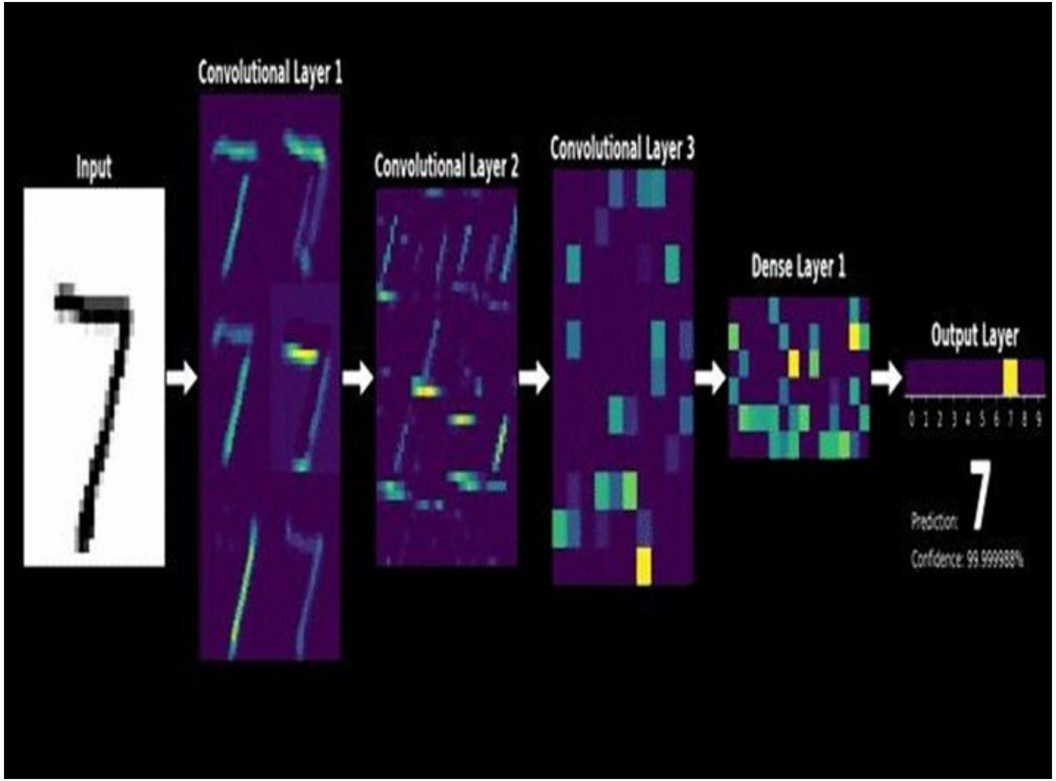
SW dataset:

- Solar wind Speed: ACE
- Location: L1 Lagrangian point
- Temporal resolution :2 hours
- Link: http://www.srl.caltech.edu/ACE/ASC/level2/



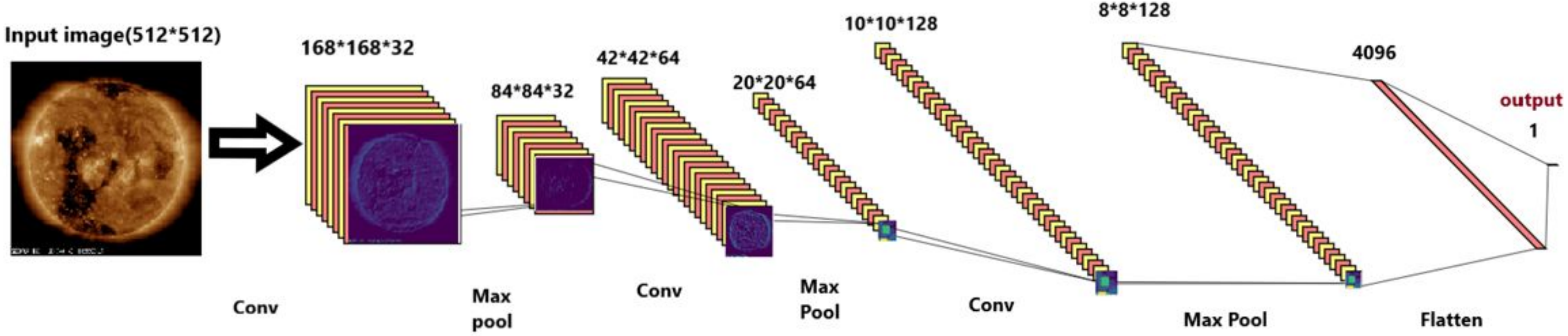
Raju and Das (2021)

Solar wind speed prediction: CNN



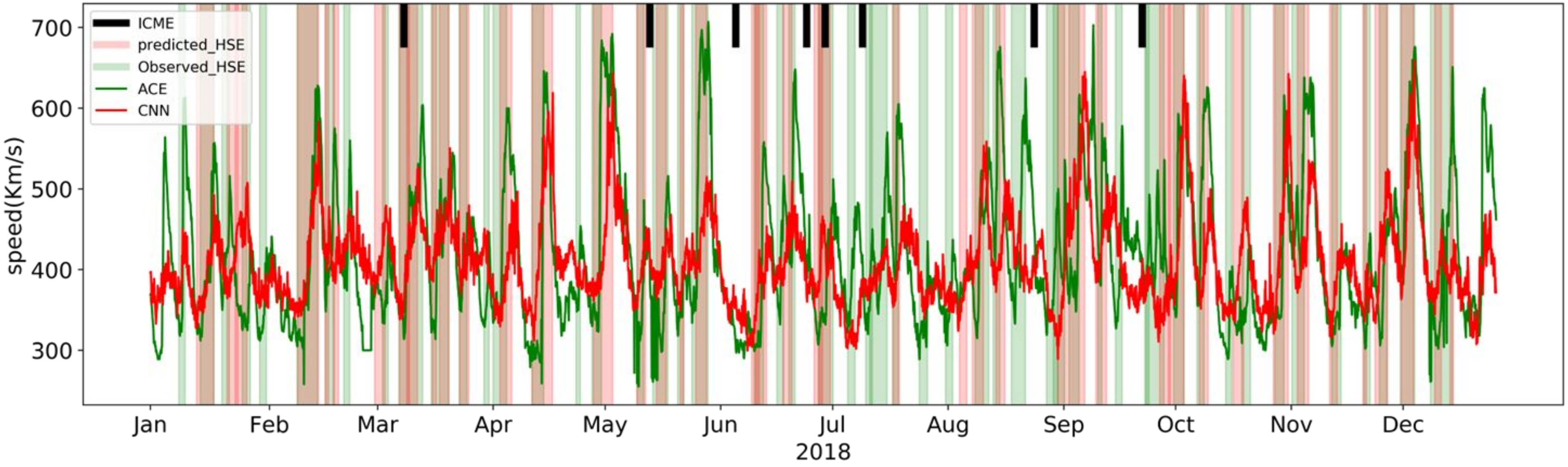
Raju and Das (2021)

Solar wind speed prediction: CNN Model



Raju and Das (2021)

Space weather Prediction : Solar wind speed (Results)



Raju and Das (2021)

Solar wind speed prediction: Visualization

Gradient Weighted Class Activation Mapping

- Used Gradcam Visualization technique with python package ELI5

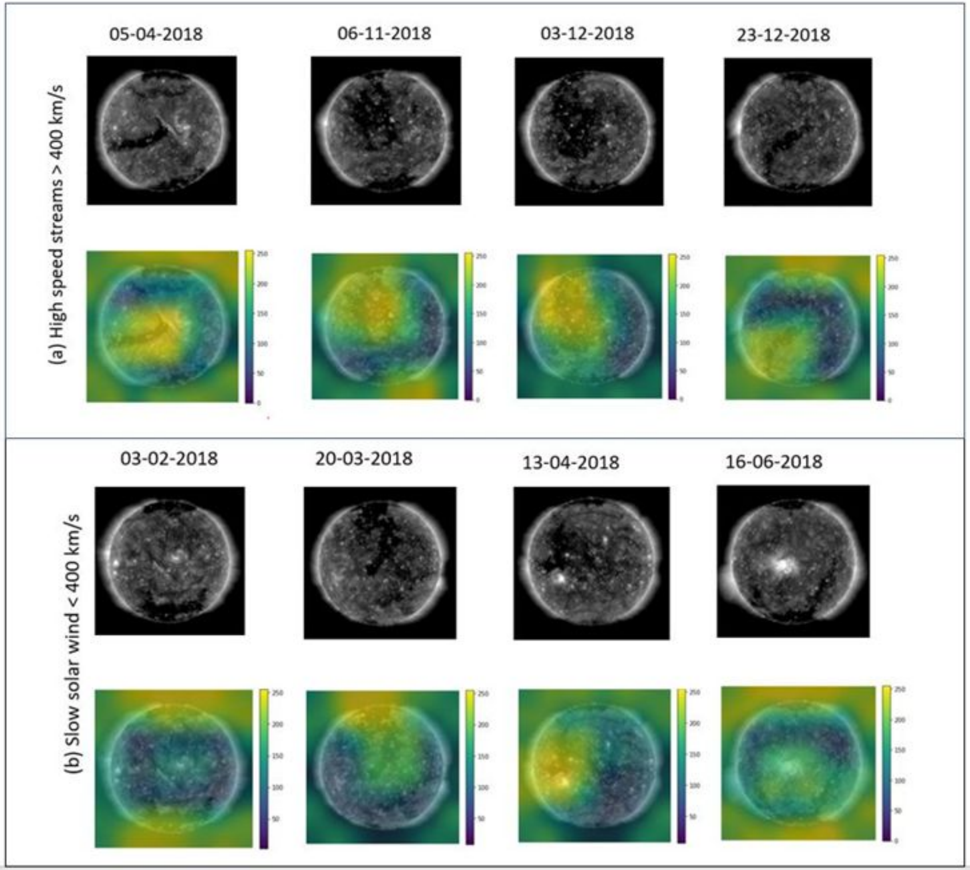
High speed streams:

- CNN activated Coronal Hole regions

Slow solar wind:

- Source of slow solar wind is still a debated topic

- CNN had activated active regions, polar CHs.

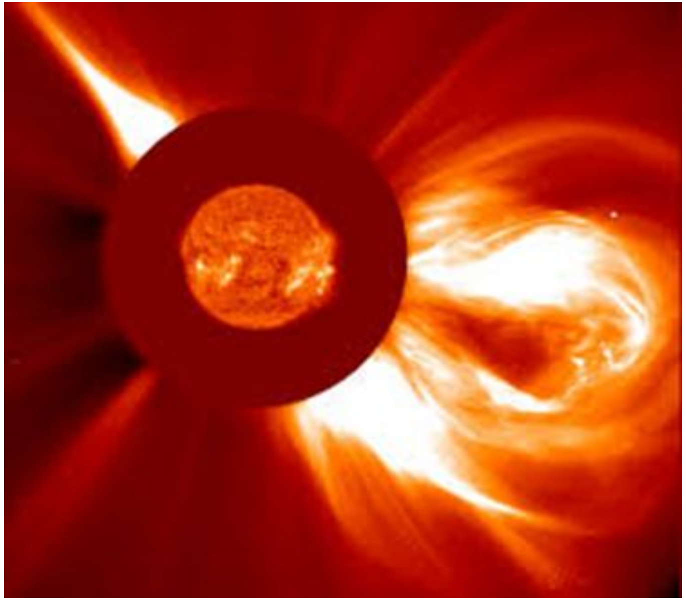


Activation heatmap of CNN predictions for high speed and slow solar wind.

Raju and Das (2021)

Predicting flare associated CMEs

Prediction of CMEs associated with Flares - Problem Statement



- More than 90% CMEs were associated with large class (M/X) flares (Bobra et al., 2015)

- Whether a flare will be associated with CME?

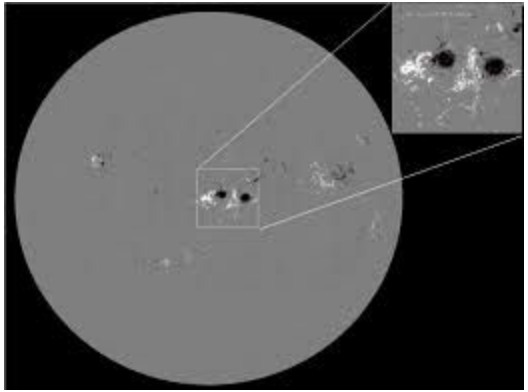
- Understanding the conditions that lead up to a CME or flare

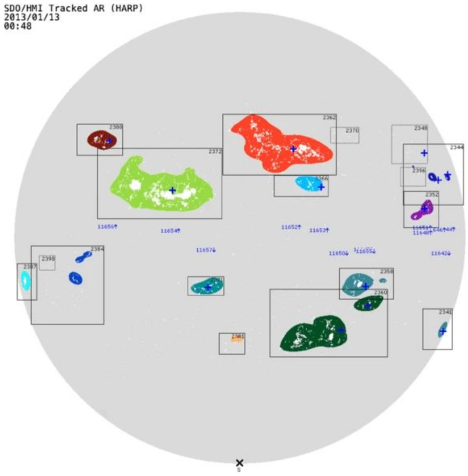
Why ML?

- There is no single parameter that clearly distinguishes between flares associated with CMEs and the flares that are not associated with CMES
- Previous statistical studies are based on individual correlation of flare parameters with occurrence of CMEs.
- Combination of parameters for huge data manually will be a tedious task

Prediction of CMEs associated with Flares - Data source

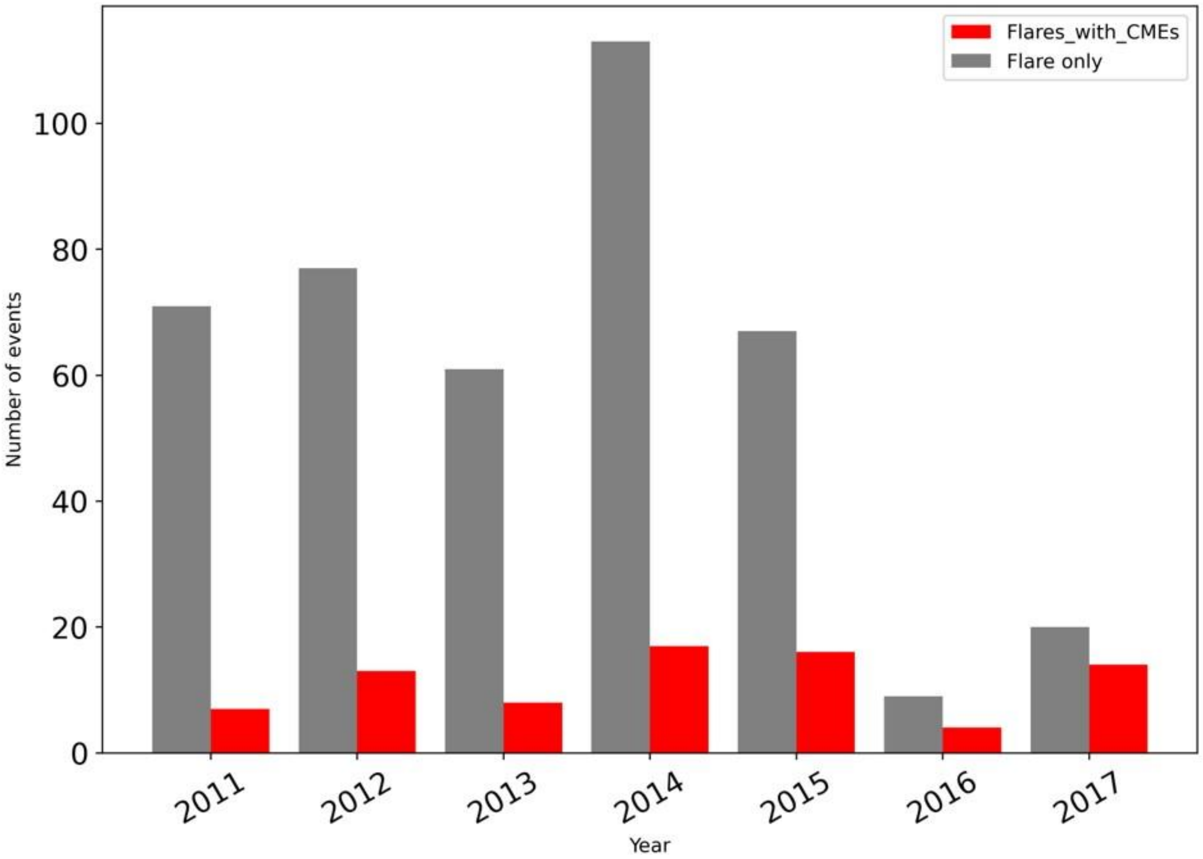
Keyword	Description	Formula
MEANGBH	Mean gradient of horizontal field	$\overline{\lvert\nabla B_h\rvert} = \frac{1}{N}\sum\sqrt{\left(\frac{\partial B_h}{\partial x}\right)^2 + \left(\frac{\partial B_h}{\partial y}\right)^2}$
MEANJZH	Mean current helicity (B_z contribution)	$\overline{H_c} \propto \frac{1}{N}\sum B_z \cdot J_z$
MEANALP	Mean characteristic twist parameter, α	$\alpha_{\text{total}} \propto \frac{\sum J_z \cdot B_z}{\sum B_z^2}$
MEANGBT	Mean gradient of total field	$\overline{\lvert\nabla B_{\text{tot}}\rvert} = \frac{1}{N}\sum\sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2}$
MEANPOT	Mean photospheric magnetic free energy	$\overline{\rho} \propto \frac{1}{N}\sum(\boldsymbol{B}^{\text{Obs}} - \boldsymbol{B}^{\text{Pot}})^2$
MEANSHR	Mean shear angle	$\overline{\Gamma} = \frac{1}{N}\sum \arccos\left(\frac{\boldsymbol{B}^{\text{Obs}} \cdot \boldsymbol{B}^{\text{Pot}}}{\lvert \boldsymbol{B}^{\text{Obs}} \rvert \lvert \boldsymbol{B}^{\text{Pot}} \rvert}\right)$
SHRGT45	Fraction of Area with Shear $>45°$	Area with Shear $>45°$/Total Area
TOTPOT	Total photospheric magnetic free energy density	$\rho_{\text{tot}} \propto \sum(\boldsymbol{B}^{\text{Obs}} - \boldsymbol{B}^{\text{Pot}})^2 dA$
MEANJZD	Mean vertical current density	$\overline{J_z} \propto \frac{1}{N}\sum\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)$
USFLUX	Total unsigned flux	$\Phi = \sum\lvert B_z\rvert dA$
MEANGAM	Mean angle of field from radial	$\overline{\gamma} = \frac{1}{N}\sum \arctan\left(\frac{B_h}{B_z}\right)$
TOTUSJZ	Total unsigned vertical current	$J_{z\text{total}} = \sum\lvert J_z\rvert dA$
ABSNJZH	Absolute value of the net current helicity	$H_{c\text{abs}} \propto \lvert\sum B_z \cdot J_z\rvert$
AREA_ACR	Area of strong field pixels in the active region	Area $=\sum$ Pixels
R_VALUE	Sum of flux near polarity inversion line	$\Phi = \sum\lvert B_{\text{LOS}}\rvert dA$ within R mask
TOTUSJH	Total unsigned current helicity	$H_{c\text{total}} \propto \sum\lvert B_z \cdot J_z\rvert$
	Flare Class	$FC = CM$
SAVNCPP	Sum of the modulus of the net current per polarity	$J_{z\text{sum}} \propto \left\lvert\sum^{B_z^+} J_z dA\right\rvert + \left\lvert\sum^{B_z^-} J_z dA\right\rvert$
MEANGBZ	Mean gradient of vertical field	$\overline{\lvert\nabla B_z\rvert} = \frac{1}{N}\sum\sqrt{\left(\frac{\partial B_z}{\partial x}\right)^2 + \left(\frac{\partial B_z}{\partial y}\right)^2}$



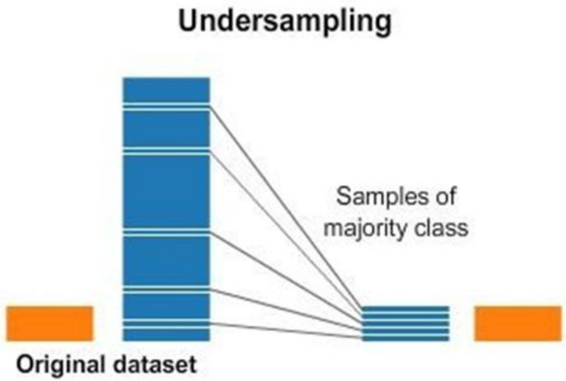


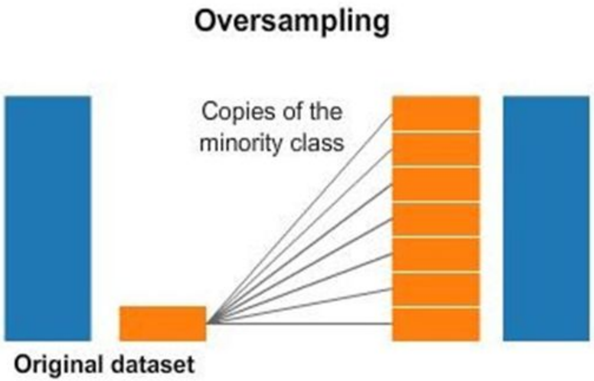
Bobra et al., 2016

Prediction of CMEs associated with Flares - Class Imbalance

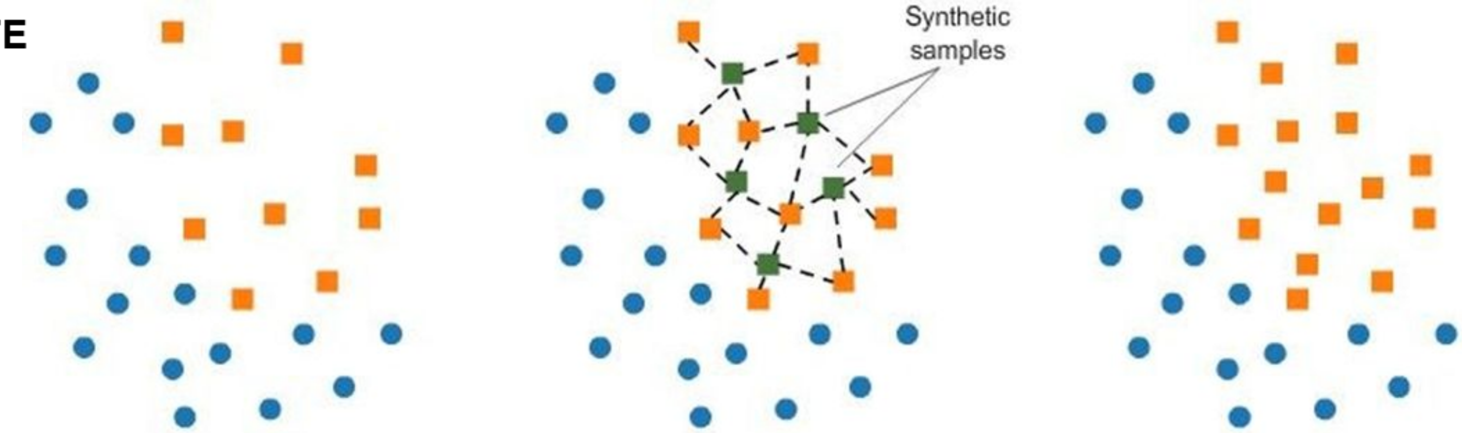


Prediction of CMEs associated with Flares - Class Imbalance

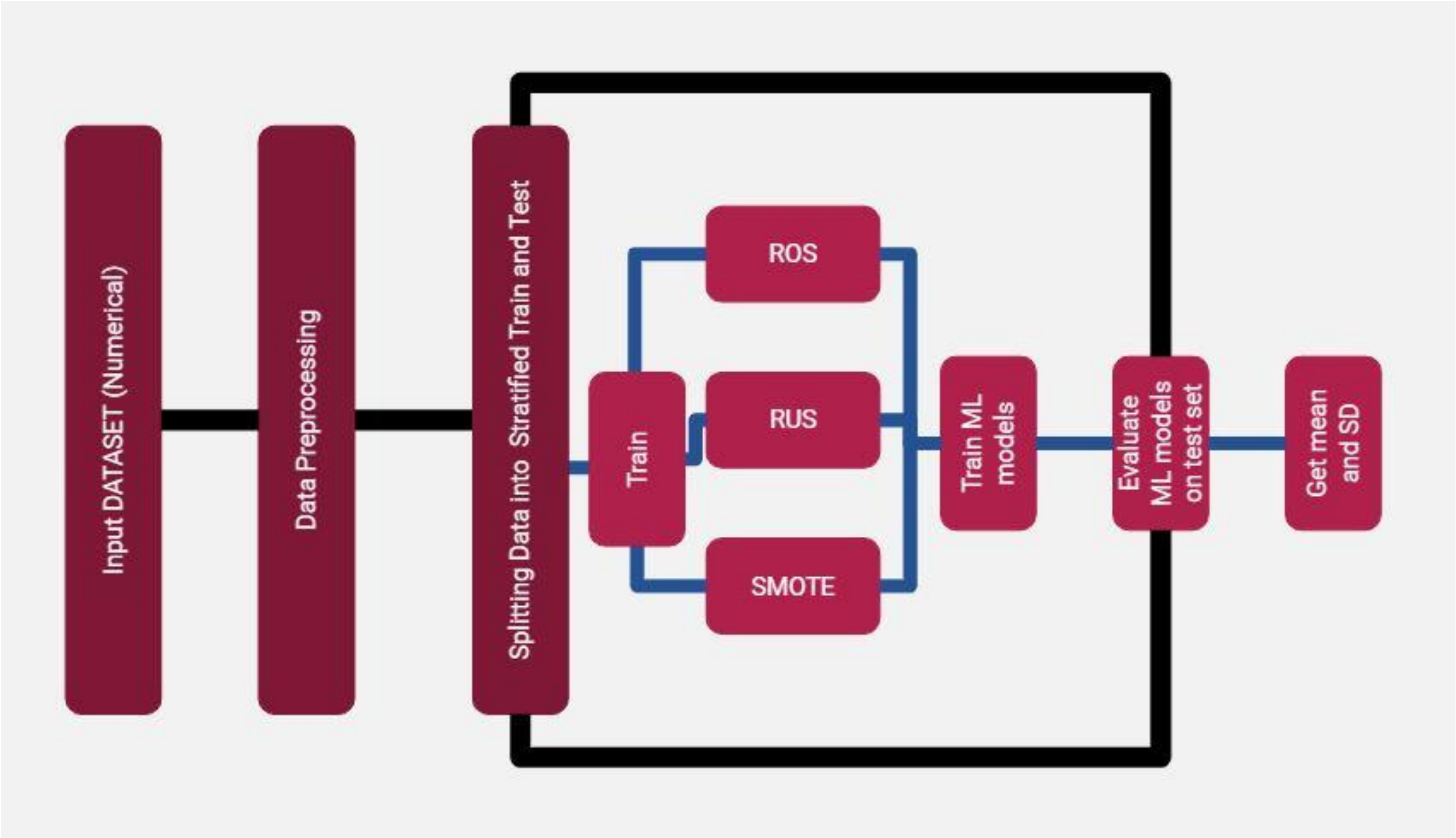




SMOTE

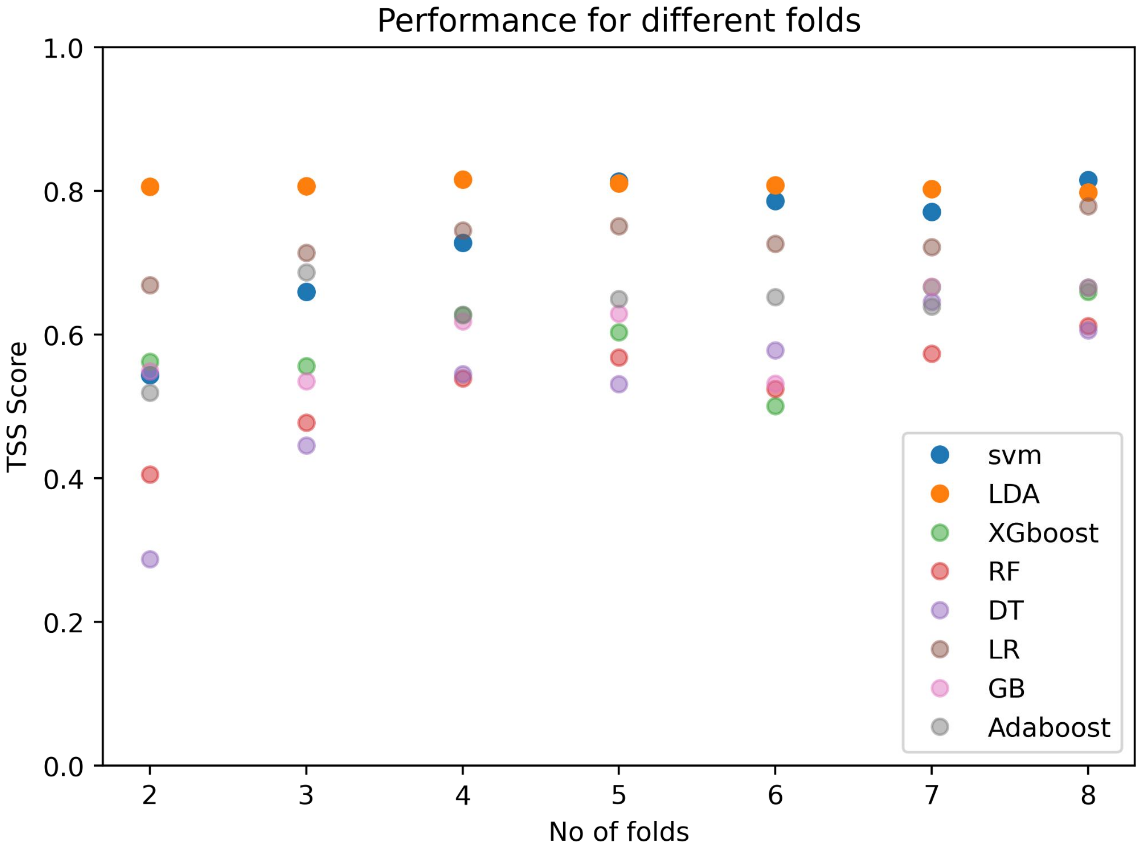


Prediction of CMEs associated with Flares - Model



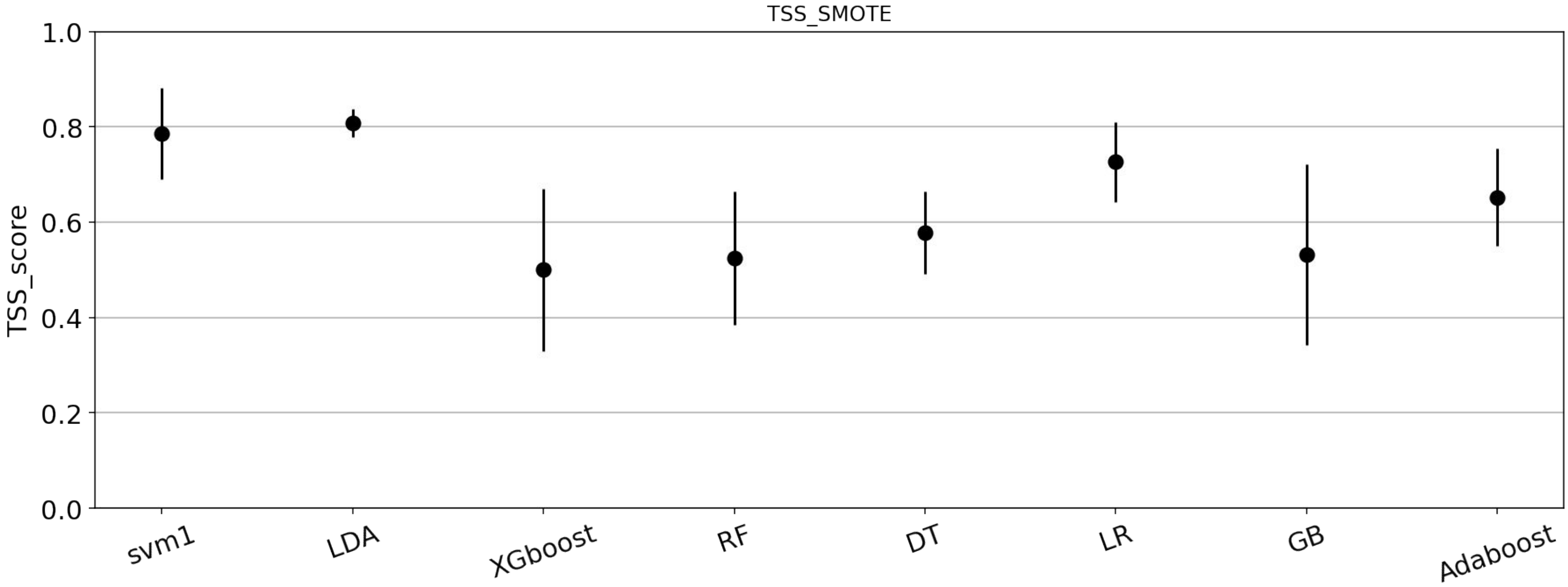
Raju and Das (in prep)

Prediction of CMEs associated with Flares - Results



Raju and Das (in prep)

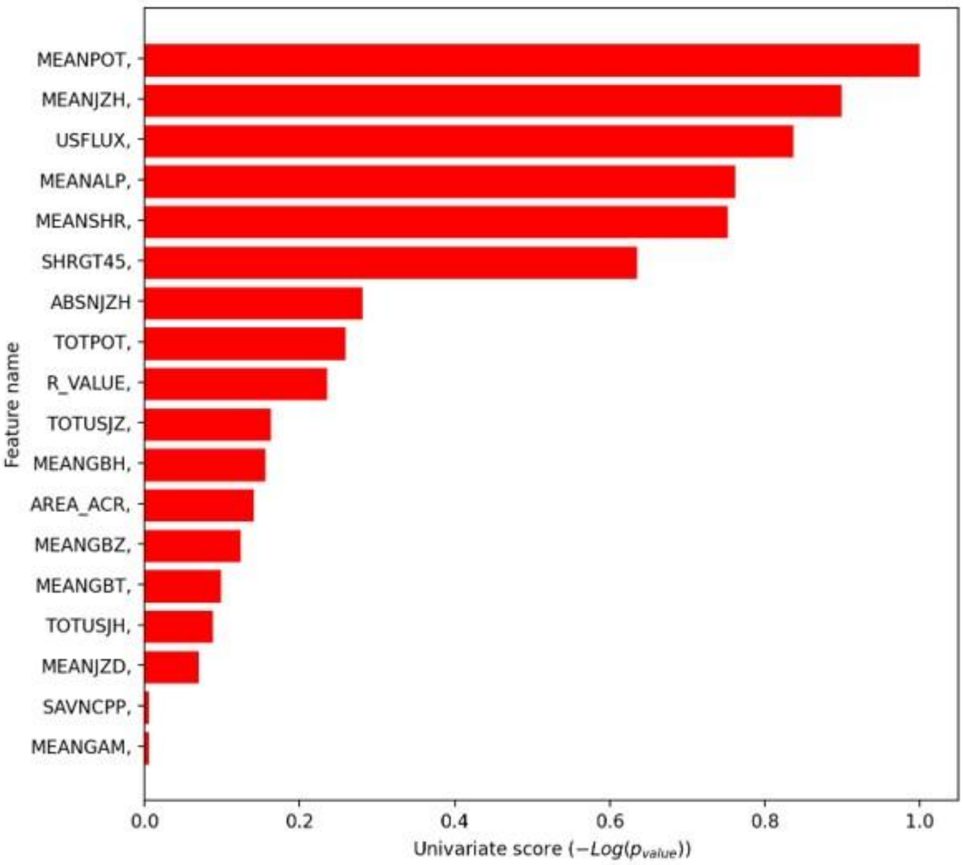
Prediction of CMEs associated with Flares - Results



$$(TSS) = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}$$

Raju and Das (in prep)

Prediction of CMEs associated with Flares - Feature importance



- evaluating the relationship between each input variable and the target variable using statistics
- Sort it based on scores

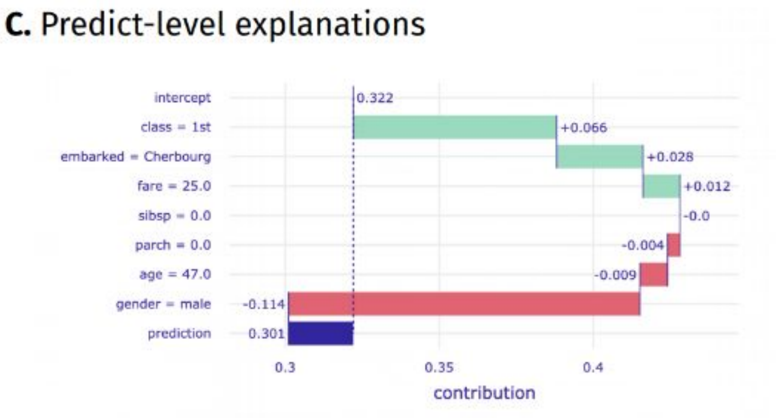
Raju and Das (in prep)

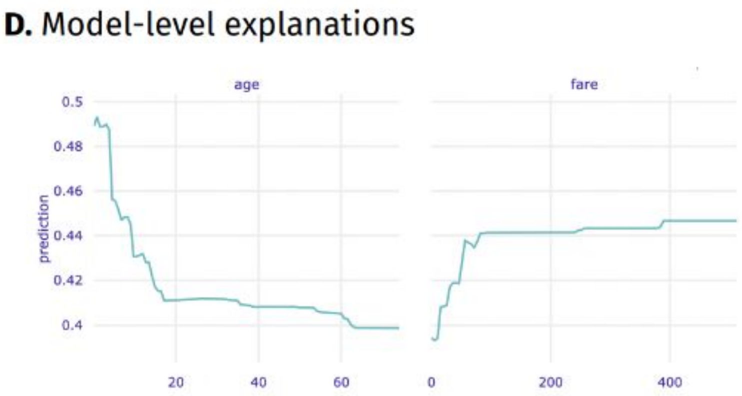
Explainable AI (Dalex) - Working

variable-importance measure

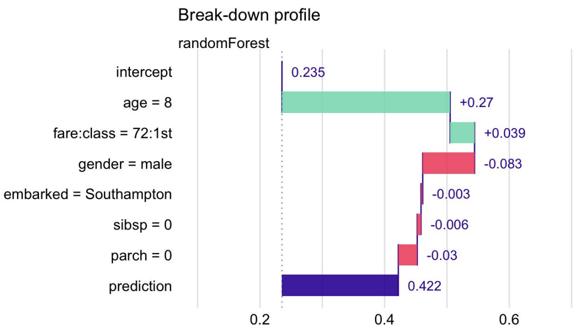
The main idea is to measure how much does a model's performance change if the effect of a selected explanatory variable, or of a group of variables, is removed?

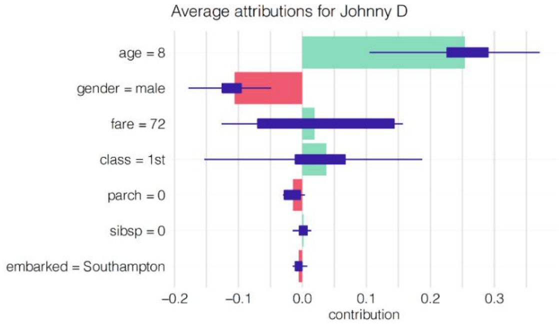
- Permute the particular variable
- calculate dropout loss
- Larger the loss, greater the importance



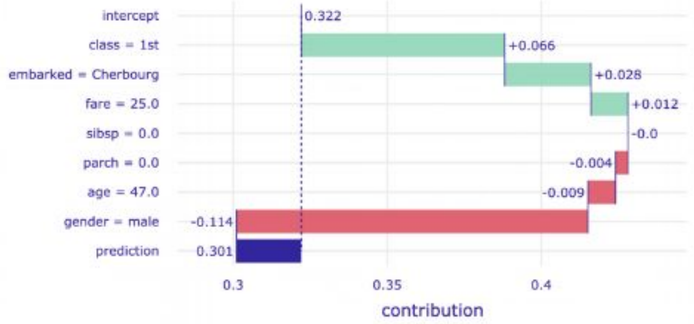


Explainable AI (Dalex) - Working

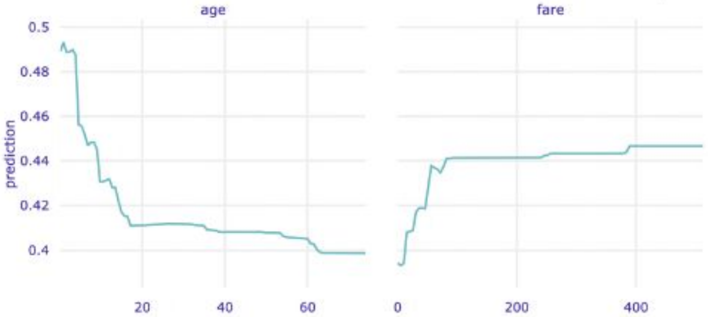




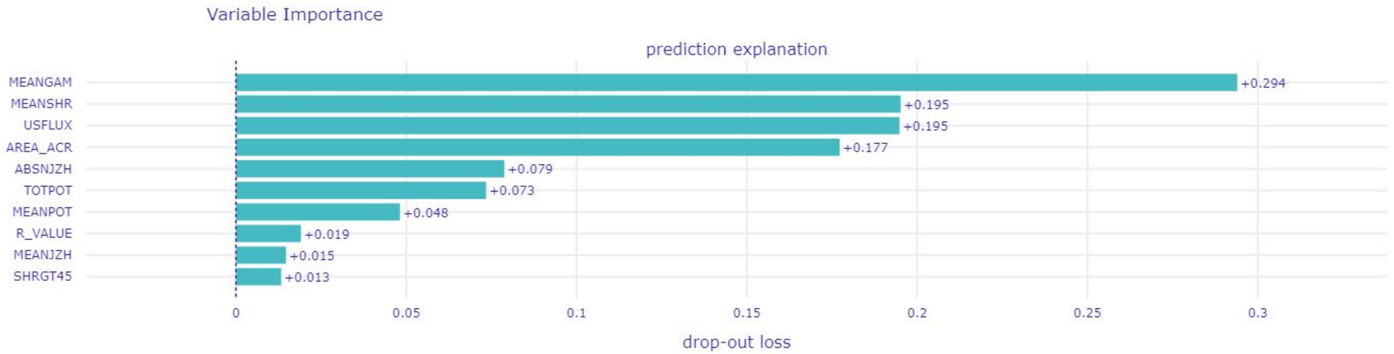
C. Predict-level explanations



D. Model-level explanations

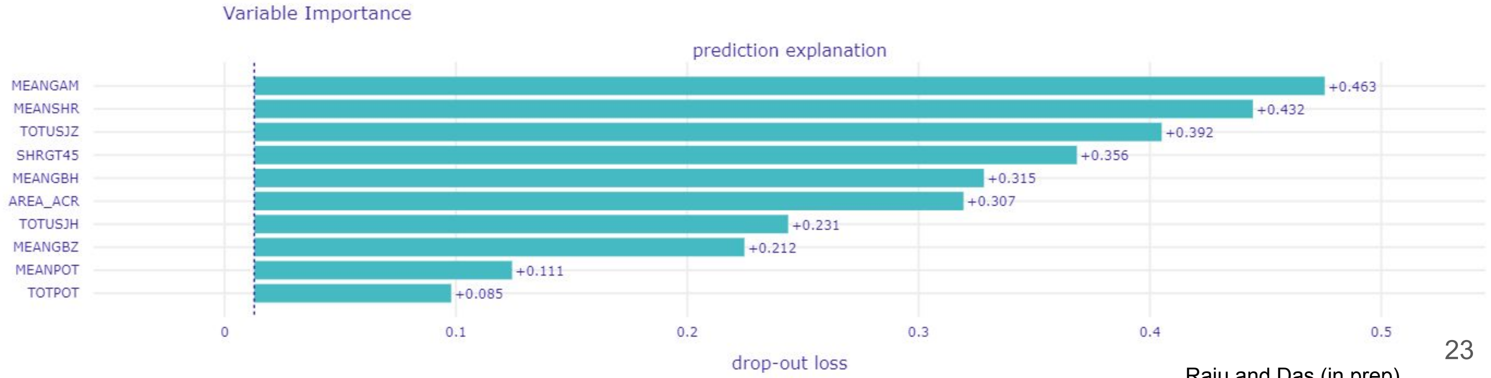


Prediction of CMEs associated with Flares - Explainable Model Predictions (Dalex)



SVM model Explanations

LDA Model Explanation



Raju and Das (in prep)

Explainable AI (Wrapper model) - Working

- **Forward selection**

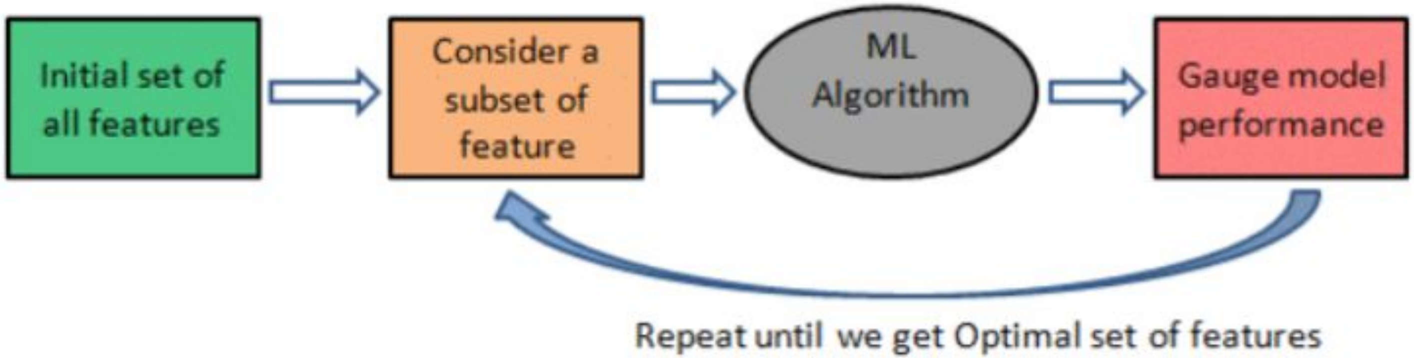
Start with null model
Add features and assess the performance

- **Backward Elimination**

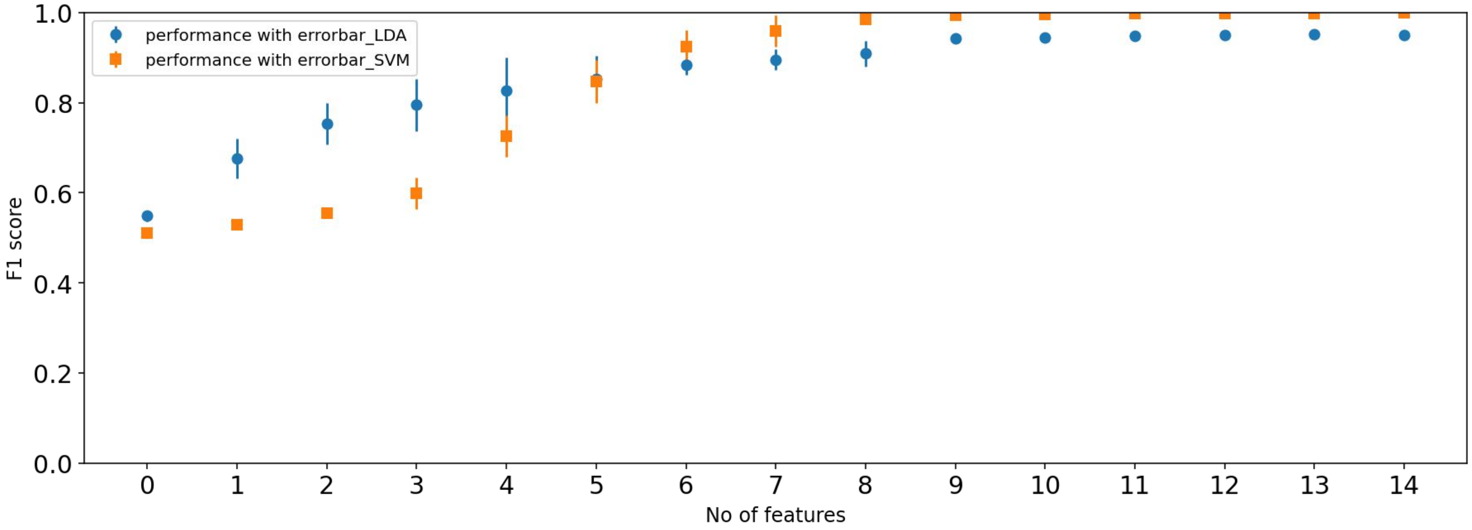
Start with full model
Eliminate features and assessing the performance

Bidirectional elimination

- Combination of forward selection and backward elimination



Prediction of CMEs associated with Flares - Explainable Model Predictions (wrapper)



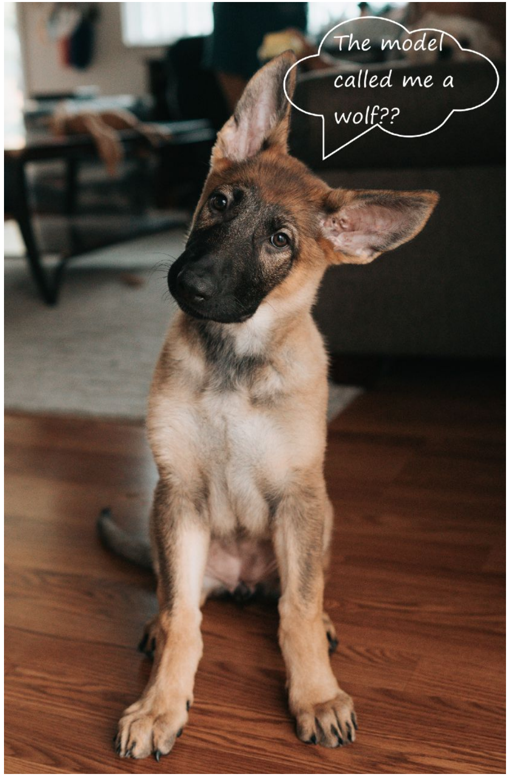
Some of the top Features_LDA: `'MEANJZH', 'MEANPOT', 'SHRGT45', 'MEANGBH', 'MEANGAM', 'MEANGBZ', 'TOTUSJZ', 'MEANSHR', 'AREA_ACR', 'R VALUE', 'ABSNJZH'`

Some of the top Features_SVM: `''USFLUX', 'MEANJZH', 'MEANPOT', 'SHRGT45', 'MEANALP', 'MEANGAM', 'SAVNCPP', 'TOTPOT', 'MEANSHR', 'AREA_ACR', 'R_VALUE'`

Raju and Das (in prep)

Summary

- Grad Cam Visualization used for understanding CNN predictions
- Class imbalance methods and metrics explored
- ML explainable model predictions methods explored



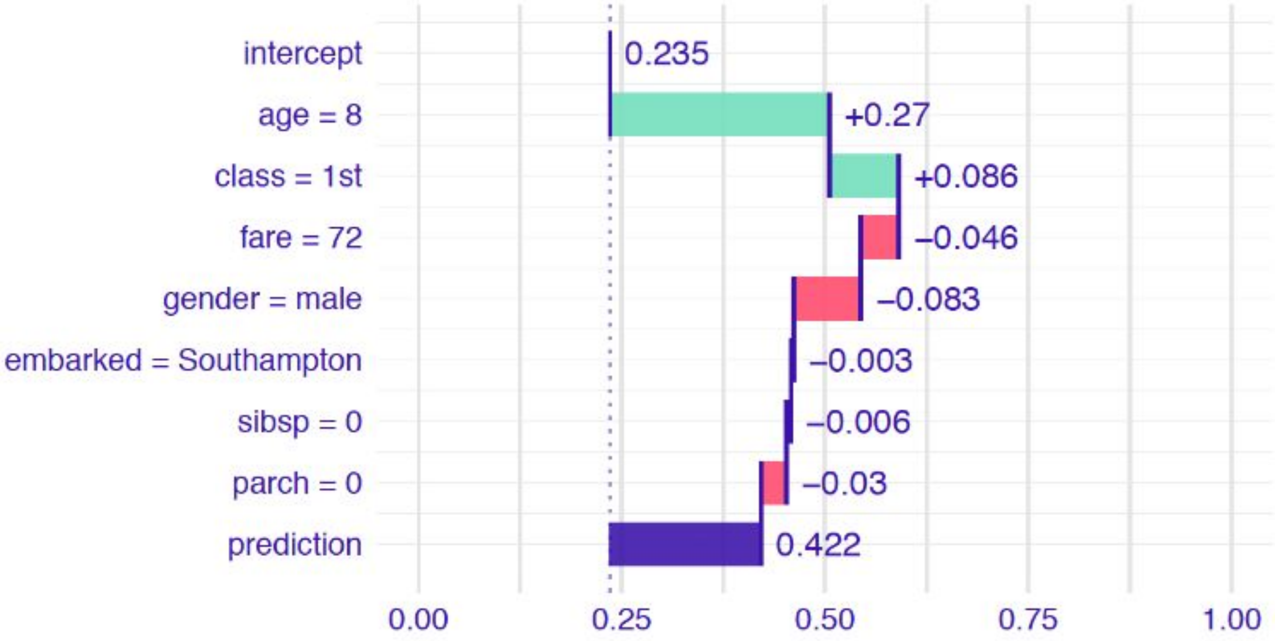

Thank you!

Title

Title

Title

C)


intercept
age = 8
class = 1st
fare = 72
gender = male
embarked = Southampton
sibsp = 0
parch = 0
prediction
0.235
+0.27
+0.086
−0.046
−0.083
−0.003
−0.006
−0.03
0.422
0.00
0.25
0.50
0.75
1.00